



Onderzoek naar hefboomen voor kwaliteitsverbetering in het onderwijs

Jaap Scheerens, Bernard Veldkamp, Martina
Meelissen

Deze rapportage bestaat uit twee delen:

Deel 1

Het rapport "Hefboomen voor onderwijsverbetering: de overheid aan zet. Wat zeggen de resultaten van onderwijs-effectiviteitsonderzoek"

Door Jaap Scheerens

Deel 2

Bijlagen bij het onderzoek naar hefboomen in kwaliteitsverbetering in het onderwijs

Door Jaap Scheerens, Bernard Veldkamp en Martina Meelissen

17 april 2021

Inhoudsopgave

DEEL 1: HEFBOMEN VOOR ONDERWIJSVERBETERING: DE OVERHEID AAN ZET	
Wat zeggen de resultaten van onderzoekseffectiviteitsonderzoek?	6
VOORWOORD.....	7
SAMENVATTING	10
1 INLEIDING	17
1.1 Vragen van de opdrachtgever	17
1.2 Het perspectief om deze vragen te beantwoorden	18
2 EEN OVERKOEPELEND CONCEPTUEEL KADER.....	21
2.1 Het basis systeemmodel om onderwijskwaliteit te typeren	21
2.2 De hiërarchische structuur	23
2.3 Alternatieve voorstelling en uitwerking van het basismodel	24
3 WAT WERKT ER OP HET MICRO-NIVEAU (LEREN EN INSTRUCTIE)	29
3.1 Overzicht van variabelen bij leren	29
3.2 Evidentie voor effectiviteit van de voornaamste categorieën van beïnvloedbare variabelen bij leren.....	31
3.3 Overzicht van variabelen bij onderwijzen	31
3.4 Een beperkt aantal theoretisch onderbouwde dimensies van effectief onderwijzen	34
3.4.1 Het belang van de leerkracht en “pedagogical content knowledge” (pck)	34
3.4.2 “Gelegenheid om te leren” (OTL) en curriculum-alignment	36
3.4.3 Gestructureerd onderwijs	39
3.4.3.1 “Directe instructie” (d.i.) als kern: didactische ondersteuning	39
3.4.3.2 Ondersteuning in het kader van klassenmanagement, leeromgeving&klimaat	41
3.4.4.. Evaluatie en feedback	45
3.5 Evidentie voor effectiviteit van de voornaamste categorieën van beïnvloedbare variabelen bij onderwijzen	46
3.6 Hoe het beleid deze kennis over effectief onderwijzen zou kunnen gebruiken	48
4 WAT WERKT ER OP HET MESO-NIVEAU (SCHOOLORGANISATIE EN SCHOOLLEIDERSCHAP)	51
4.1 Overzicht van variabelen geadresseerd in school effectiviteitsonderzoek	51
4.2 Theorieën over de school als organisatie	54
4.3 Empirische resultaten van schooleffectiviteitsonderzoek.....	64
4.4 Discussie en aangrijpingspunten voor beleid	68
5 WAT WERKT OP SYSTEEMNIVEAU?.....	71
5.1 Inleiding	71
5.2 Overzicht van variabelen op systeemniveau	71
5.2.1 Een brede beschrijving van systeemindicatoren	71
5.2.2 Een beperkter aantal kernfactoren	74

5.3	Conceptualisering en theoretische onderbouwing van kernfactoren	77
5.3.1.	Inleiding, verschillende interpretaties van het rationaliteitsparadigma	77
5.3.2.	Keuze-vrijheid en functionele decentralisatie	77
5.3.3	Evaluatie en Accountability	79
5.3.4	De stratificatie van scholenstelsels; geïntegreerd versus categoriaal	80
5.3.5	Facetten van “lopend” beleid	80
5.4	Resultaten van empirisch onderzoek	82
5.4.1	Illustratieve onderzoeksresultaten	83
5.4.1.1	Keuzevrijheid, privé-scholen en autonomie	83
5.4.1.2	Accountability	85
5.4.2	Aanvullende resultaten op basis van recente PISA-onderzoeken	87
5.4.2.1	Samenvatting van de effecten van systeemvariabelen in PISA ¹	87
5.4.2.2	Secundaire analyses op PISA 2009 data	89
5.4.3	Impressies van enkele facetten van “lopend beleid” en hypothesen over de effecten ervan	91
5.4.3.1	Het onderwijsbudget	91
5.4.3.2	Nationale beleidsprogramma’s gericht op kwaliteitsverbetering	92
5.4.3.3	Initiatieven tot curriculumvernieuwing	93
5.4.3.4	Beleid gericht op het tegengaan van achterstanden en ongelijkheid in het onderwijs	93
5.4.3.5	De opleiding van leraren	93
5.4.3.6	Samenvattende impressie van de facetten van lopend beleid	95
5.5	Conclusies	95
6	HEFBOMEN ALS SCENARIO’S VOOR VERBETERING EN HET TEGENGAAN VAN INEFFECTIVITEIT ...	98
6.1	Inleiding	98
6.2	Samenvatting van empirische resultaten	98
6.3	Verklarende mechanismen voor effectiviteit en ineffectiviteit	102
6.4	“Progressively worse” (Peal, 2014)	102
6.5	Over vernieuwing en verbetering; het besturend vermogen van de overheid	105
6.6	Scenario’s	108
6.7	Suggesties voor vervolgactiviteiten	110
Bijlage 1	Literatuur Referenties	111
DEEL 2. BIJLAGEN BIJ HET RAPPORT NAAR HEFBOMEN VOOR ONDERWIJSVERBETERING	125	
ANNEX 1 Results from PISA	126	
ANNEX 2 Results from TIMSS and other international studies	134	
ANNEX 3 De rol van technologie en AI in het onderwijsproces	143	
Cv’s project team	145	

¹ Over dit onderdeel is een uitvoeriger ANNEX toegevoegd, waarin de resultaten meer in detail zijn weergegeven.

DEEL 1: HEFBOMEN VOOR ONDERWIJSVERBETERING: DE OVERHEID AAN ZET

**Wat zeggen de resultaten van
onderwijseffectiviteitsonderzoek?**

Door Jaap Scheerens

VOORWOORD

In dit rapport gaat het over mogelijke aangrijpingspunten voor overheidsbeleid om bij te dragen aan de verbetering van het onderwijs, voor zoverre deze gebaseerd kunnen worden op bestaande kennis over onderwijs-effectiviteit. In dit voorwoord zou ik willen wijzen op enkele beperkingen en nuances. Dat dit voorwoord hiermee enigszins het karakter van een “disclaimer” heeft, neemt niet weg dat ik ervan uitga dat de synthese van kennis en onderzoeksresultaten relevante informatie aan het beleid kan leveren over wat wel, en ook over wat niet te doen.

Tussen bewijskracht en speculatie

De kernvragen bij onderwijs-effectiviteitsonderzoek zijn “wat werkt”, welke beïnvloedbare factoren hebben invloed op onderwijsprestaties en de vervolgvraag “waarom” die factoren werken. Met kennis over onderwijs-effectiviteit worden zowel resultaten van empirisch onderzoek (wat werkt?) bedoeld als gefundeerde interpretaties daarvan (waarom werkt het?). In het beste geval zijn deze interpretaties gebaseerd op gevestigde wetenschappelijke theorie, maar ook op ondersteuning door overige, contextuele gegevens en in andere gevallen zelfs alleen op basis van plausibele redeneringen. De status en de zeggingskracht van de gegevens variëren met de mate van onderbouwing en de grootte van effecten, maar hangen ook samen met de theoretische interpretatie, de ondersteuning door overige empirische gegevens en de plausibiliteit van hypothetische redeneringen.

Ter illustratie zou ik willen verwijzen naar een van de best gedocumenteerde uitkomsten uit internationaal vergelijkend assessment onderzoek (PISA, TIMSS). Het gaat om de bevinding dat categorale scholenstelsels in het voortgezet onderwijs en schoolkeuze op betrekkelijk jonge leeftijd, meer ongelijkheid en soms ook lagere prestaties laten zien dan geïntegreerde scholenstelsels (*comprehensive schools*). Regelmatig wordt op basis van deze uitkomst gepleit voor het veranderen van de structuur van het Nederlandse stelsel voor voortgezet onderwijs, zeg maar, een tweede kans voor het eerder mislukte Middenschool Experiment. De OECD wijst er in het rapport over PISA 2015 echter op dat de veronderstelde negatieve effecten van een sterk gestratificeerd stelsel gemitigeerd worden door andere kenmerken van de Nederlandse context, zoals de bijna universele deelname aan vroegschoolse educatie vanaf het vierde levensjaar, de gunstig geachte combinatie van autonomie en accountability, een bovengemiddeld budget voor voortgezet onderwijs, waardering voor het leraarsberoep en de mogelijkheid van horizontale doorstroming. Ook wijst men erop dat Nederland op ongelijkheidsindicatoren vergelijkbaar scoort met het OECD-gemiddelde en het in dit opzicht beter doet dan andere landen met een sterk gestratificeerd stelsel. Mede gezien de historie in het denken over stelselwijziging (vergelijk het rapport van de Commissie Dijsselbloem) en de hoge transactiekosten zou de suggestie aan de overheid kunnen zijn om deze hefboom maar liever ongebruikt te laten.

De focus op “basiskwaliteit”

De uitkomsten van onderzoek die worden besproken in dit rapport hebben betrekking op leerprestaties in basisvakken, taal, lezen, rekenen/wiskunde en natuurwetenschappen. In het recente onderzoek van McKinsey & Company naar de doelmatigheid en de toereikendheid van het funderend onderwijs wordt in de diagnose over het Nederlandse onderwijs de noodklok geluid over deze basiskwaliteit, om vervolgens te concluderen dat ze “op orde” is, maar dat er wel miljarden moeten worden besteed aan ondersteuning van scholen om te voldoen aan “additionele verwachtingen” en “schooleigen ambities”. Hier lijkt iets mis in de aansluiting tussen probleemstelling en oplossingen. Op meerdere plaatsen in dit

rapport wordt de vraag opgeroepen naar de prioriteiten bij vernieuwings- en verbeteringsactiviteiten, “smalle” en brede vorming, en aandacht voor het primaire onderwijsproces of secundaire functies van scholen. Deze normatieve keuzes staan niet centraal, maar kunnen evenmin vermeden worden. Het uitgangspunt in dit rapport is dat basisvakken nog steeds de kern van het onderwijs vormen.

Methodologische beperkingen

In dit rapport wordt de effectiviteit van maatregelen, processen en interventies op drie niveaus behandeld, het microniveau (leren en instructie), het mesoniveau (de school als organisatie) en het macroniveau (structurele kenmerken en maatregelen op stelselniveau). Binnen de context van dit rapport is het niet aan de orde om op de vele methodologische aspecten in te gaan. Wel is het van belang op te merken dat het onderzoek op het microniveau onder beter gecontroleerde condities kan plaatsvinden dan het geval is met het onderzoek op school- en systeemniveau (Zie Scheerens, 2016, 2017) en daarom door de bank genomen van betere kwaliteit is.

Over hefboomen en maakbaarheid

Een van de grote uitdagingen van onderwijs-effectiviteitsonderzoek is het isoleren van de invloed van interventies en beheersbare processen enerzijds en gegeven omgevingscondities, die ook invloed hebben op de uitkomsten, anderzijds. Meestal blijkt dat naarmate men hier beter in slaagt de effecten waarin men geïnteresseerd is kleiner zijn. Ook is er meer variantie tussen leerlingen dan tussen leraren, scholen en nationale onderwijsstelsels. Zo bleek bijvoorbeeld uit gegevens van PISA 2015 dat het percentage totale variantie in leerling-prestaties in natuurwetenschappen tussen landen 10% bedroeg, tussen scholen 28% en tussen leerlingen 62%. Uit schooleffectiviteitsonderzoek blijkt dat er meestal grotere lerareneffecten dan schooleffecten worden gevonden. Zo vonden Kyriakides en Luyten (2009) bijvoorbeeld 13% van de totale variantie verklaard door scholen tegen 17.7% verklaard door leraren. Wanneer dit soort “bruto” effecten wordt gecorrigeerd voor achtergrondkenmerken van leerlingen, zoals eerdere prestaties, worden de “netto-effecten” nog kleiner, voor schooleffecten in de orde van 4 tot 10% (Scheerens, 2017). In het algemeen gaat het dus om kleine tot zeer kleine effecten. Ook uit andersoortige gegevens kan men opmaken dat de maakbaarheid van onderwijs in de zin van verbetering van prestaties een taaie materie is. Tijdreeksgegevens op basis van PISA uitkomsten laten een verbluffende stabiliteit in prestatieniveaus van landen zien (Aloisi & Tymms, 2017). Ook geldt dat de interventies mogelijk vaak te zwak zijn om effect te sorteren. Het effect van klassenverkleining is hiervan een voorbeeld, nauwelijks effect van een verkleining met 2 of 3 leerlingen, wel een effect als het gaat om een (zeer kostbare) verkleining met 10 leerlingen. Een (enigszins) positieve keerzijde van deze medaille is dat, door de bank genomen, goed scorende stelsels en scholen moeilijk uit hun evenwicht geraken. Een Nederlands voorbeeld is dat de zwaar bekritiseerde vernieuwing van de 2efase en het “Studiehuis” de goede prestaties van Nederlandse scholen op de middellange termijn niet noemenswaardig hebben doen dalen². Een relativerende interpretatie van “kleine effecten” van kwaliteitsverschillen tussen leerkrachten en scholen die wel wordt gegeven, is dat een goede school of leerkracht een multipliereffect heeft doordat grote aantallen leerlingen ervan profiteren. Deze redenering zou zeker ook opgaan voor maatregelen op stelselniveau. Als we een goede hefboom voor overheidsbeleid zouden kunnen vinden en implementeren, is dat meteen ook een flinke knuppel.

² Tenzij bepaalde tegenvallende resultaten van de laatste 5 jaar het gevolg zijn van een “sleeper effect”

Ineffectiviteit

Kennis over eventueel plausible investeringen en maatregelen die *niet* effectief zijn, is minstens zo belangrijk als kennis over effectieve praktijken. Dit is een andere manier om de resultaten van onderwijs-effectiviteitsonderzoek te bekijken. In het rapport te bespreken productiviteitsanalyses laten een negatieve productiviteitslijn zien voor het funderend onderwijs (Blank & Van Heezik, IPSE, 2019). Tegelijkertijd is er zeer veel activiteit rondom nieuwe onderwijsconcepten en uitgebreide ontwikkeling van netwerken en instanties die zich met begeleiding en professionalisering bezighouden. In de internationale literatuur is een kritische benadering ontstaan van “progressief” onderwijs dat sterk overeenkomt met het vroegere Studiehuis en het huidige gepersonaliseerd leren. Zonder ons al te veel te begeven in politieke kwesties is het interessant te kijken naar wetenschappelijke kritiek op de achtergrond van deze bewegingen.

De uitdaging van kennisbenutting en implementatie

De boodschap vanuit onderwijs- effectiviteitsonderzoek aan het beleid en de onderwijspraktijk over de laatste 30 jaar is tamelijk consistent. Op de dag dat ik dit schrijf vind ik het volgende commentaar in de Volkskrant;

“Het is de tragiek van het Nederlandse onderwijs: we weten en meten alles, we hebben uitstekende onderzoekers en didactici en kennen de oorzaken van mislukking. Maar al die kennis bereikt zelden de praktijk; de kinderen profiteren er nauwelijks van. Als we daar nu eens de komende jaren op inzetten” (Column van Aleid Truijens, VK 9 maart, 2021). Een prikkelende uitspraak waarover een *reality check* desalniettemin nodig is. Maar op zijn minst enkele hypotheses over hoe dit zou kunnen komen zullen in dit rapport niet ontbreken.

Ik ben dankt verschuldigd aan Eef Jacobs en AnneLuc van der Vegt (Oberon), die eerdere versies van dit rapport van kritisch commentaar hebben voorzien. Marja Knecht heeft zorg gedragen voor de lay-out van het rapport. De aanbevelingen tot verder onderzoek in Hoofdstuk 6, par. 6.7, zijn geformuleerd in samenspraak met Bernard Veldkamp en Martina Meelissen van de UT.

Jaap Scheerens

SAMENVATTING

Dit rapport is gebaseerd op de gedachte dat de resultaten van empirisch onderzoek naar onderwijseffectiviteit van betekenis zijn voor een overheid die moet zorgen voor de handhaving en zo mogelijk verbetering van de onderwijskwaliteit. Uiteraard is er geprobeerd om deze verwachting waar te maken.

Maar de lat ligt hoog en er zijn wel beperkingen. Ook deze worden in het rapport toegelicht.

De kennis over onderwijseffectiviteit op systeem- school- en klasniveau wordt weergegeven in gedetailleerde overzichten van specifieke variabelen, maar ook op meer samenvattende dimensies. Een overzicht van deze dimensies is gegeven in de onderstaande figuur.

Tabel S1 Effectiviteitsdimensies op micro-, meso- en macroniveau

Effectiviteitsdimensies micro-niveau
Goed opgeleide leerkrachten
Genoeg tijd en gelegenheid voor leren in basisvakken
Gestructureerd onderwijs (didactiek)
Gestructureerd onderwijs (management en klimaat)
Goede voorzieningen voor evaluatie en feedback
Effectiviteitsdimensies meso-niveau
Een schoolorganisatie die effectieve instructie ondersteunt
Een schoolorganisatie die onderhevig is aan output-gerichte accountability, en die deze gebruikt om een prestatiegerichte focus te hebben
Een schoolorganisatie die goed gebruik maakt van evaluatie en feedback-processen
Een schoolorganisatie die efficiënt is in leiderschapsfuncties (onderwijskundig leiderschap) en zuinig is met investeringen op secundaire organisatie factoren
Een schoolorganisatie die gebruikt maakt van formele regels om een ordelijk klimaat te stimuleren
Effectiviteitsdimensies macro-niveau
Aantrekkelijke arbeidsvoorwaarden voor leerkrachten, actief recruteringsbeleid.
Een kwalitatief hoogstaande lerarenopleiding. Bevordering van een veelzijdig en goed geïntegreerd onderwijs repertoire bij aanstaande leerkrachten ("pedagogical content knowledge")
Een goed uitgelijnd curriculum

Tabel S2 Besturingsmechanismen macro-niveau

Besturingsmechanismen macro-niveau
Synoptische planning
Marktwerking/autonomie
Evaluatie gestuurde planning
New public management
Stelselherziening (adaptieve planning)

In het rapport worden deze dimensies beschreven en toegelicht en verder voorzien van theoretische onderbouwing. Daarnaast worden de resultaten van empirisch onderzoek beschreven en samengevat. Waar dat kan gebeurt dit in de vorm van "effectgroottes" die de sterkte van het verband tussen de effectiviteitsbevorderende factoren en leerprestaties uitdrukken. In de volgende overzichtstabel worden deze resultaten samengevat, met vermelding van theoretische basis en ineffektieve tegenhangers.

Tabel S3 Effectiviteitsbevorderende factoren op micro-niveau

Factoren micro-niveau (onderwijzen)			
Typering	Effectgrootte	Conceptuele basis	Ineffectieve tegenhanger
"Pedagogical content knowledge"	-	Didactische analyse	Relativering van het belang van kennis
Gelegenheid om te leren (OTL)	.30	Een goed uitgelijnd curriculum	De mythe van inhoudsloze vaardigheden
Klassenmanagement van effectieve leertijd	.20	Carroll model	
Directe instructie	.50	Cognitieve informatieverwerkingstheorie	"Open", ongeleide instructie, geïnspireerd door het constructivisme
Vakgebonden leerstrategieën en meta-cognitie	.40	Doelstellingstaxonomieën Cognitieve informatieverwerkingstheorie	
Ordelijk klimaat	.30	Formele regulering; theorie over de bureaucratie	Ontbreken van regels en orde "over-permissive"
Evaluatie en feedback	.20	Cybernetisch principe	Anti-toets lobby

Tabel S4 Effectiviteitsbevorderende factoren op mesoniveau

Factoren meso-niveau (school beleid en organisatie)			
Typering	Effectgrootte	Conceptuele basis	Ineffectieve tegenhanger
Schoolvariabelen die staan voor aggregaties van condities op microniveau (OTL, schoolcurriculum, leertijd, huiswerk en prestatiegerichtheid)	.30	Professionele bureaucratie; leerkracht autonomie Rationeel geleide school	Autocratisch leiderschap, laissez faire beleid
Schoolvariabelen die uitdrukking geven aan directe ondersteuning van het primaire proces. Instructiegericht leiderschap, duidelijke schoolregels en taakgerichte samenwerking	.20	Model van onderwijskundig leiderschap en de lerende organisatie, waarin evaluatie en feedback het centrale mechanisme zijn	Accent op secundaire condities en niet essentiële taakaspecten van leraren (zoals onderzoek, en curriculumontwikkeling)
Algemene organisatiefactoren die gericht zijn op het welzijn van leerkrachten en het onderhouden van externe relaties; transactioneel leiderschap	.10	De school als open systeem en als sociale werkgemeenschap.	Overaccentuering die ten koste gaat van de aandacht voor het primaire proces.

Tabel S5 Effectiviteitsbevorderende factoren op macro-niveau

Factoren macroniveau (overheidsbeleid)			
Typering	Indruk van effect	Conceptuele basis	Ineffectieve tegenhanger
Vrije schoolkeuze & privatisering	Geen effect	Economische theorie, marktwerking	
Autonomie	Geen tot zeer klein effect	Minder regeldruk, meer responsiviteit	
Accountability	Klein effect	<i>Public choice</i> theorie; prestatieprikkel; theorie over curriculum-alignment	Immunisering tegen externe evaluatie
“New public management”	Klein effect	Functionele decentralisatie	Geen of “zachte” opbrengstmetingen. Anti-opbrengstgerichtheid en verheffing van middelen en processen tot doelen.
Stratificatie	Betekenisvol negatief effect	Te vroege selectie; onvoldoende mogelijkheden tot doorstroming tussen school categorieën	In dit geval is er een effectieve tegenhanger, namelijk meer geïntegreerde schoolstelsels
Verhoging onderwijsbudget	Alleen effectief in arme landen	Effect van betere personele en materiële voorzieningen	Verminderende meeropbrengst en ondoelmatigheid van bestedingen
Opleidingsniveau en professionele ontwikkeling leraren	Geen effecten in OECD- landen	Verwacht positief effect van beter opgeleide leraren	
Lerarentekort	Klein negatief effect	Groot belang van voldoende personeelsbezetting	
Curriculum alignment	Klein effect van OTL	Positieve verwachtingen over verticale en horizontale “uitlijning” van het curriculum	Losse koppeling en autonomie in het curriculum domein

Wat werkt?

Microniveau – Deze uitkomsten laten zien dat er een stevig empirisch bewijs is en dat er een goede theoretische onderbouwing bestaat voor de effectiviteit van directe instructie. Relevante optimaliserende condities zijn: de erkenning van het belang van vakgerichte kennis, didactisch goed geschoolde leerkrachten, formatieve en summatieve toetsing en feedback.

Mesoniveau – De belangrijkste taak van de school is om deze optimaliserende condities zo goed mogelijk te waarborgen. Dit komt tot uitdrukking in voldoende leertijd en “opportunity” om vakgerichte kennis te verwerven en een stijl van leiding geven die de autonomie van leerkrachten respecteert, maar ook taakgericht ondersteunt.

Macroniveau – De taak van de overheid is om de kwaliteit van het onderwijs te stimuleren en te waarborgen, met name door duidelijke onderwijsdoelstellingen te formuleren, zorg te dragen voor een goed uitgelijnd systeem van examens en toetsen en te zorgen voor inspectie waarbij scholen verantwoording afleggen over hun prestaties. *Last but not least* ook het zorgen voor een lerarenopleiding van hoog niveau, waarin leraren vooral worden opgeleid om les te geven en een zekere regulering op de marktwerking bij dienstverlening en ondersteuning.

Wat werkt niet?

De voornaamste bron van *ineffectiviteit* op het microniveau is onvoldoende gestructureerd “open” onderwijs, waarbij te snel een beroep wordt gedaan op zelfregulatie. Op het mesoniveau zijn dat gebrek aan taakgerichte focus, overaccentuering van secundaire organisatie-condities en niet-essentiële functies van leraren, zoals het doen van onderzoek en curriculumontwikkeling. Problemen op het macroniveau zijn erosie van *checks & balances* bij het functioneren van het georganiseerde onderwijsveld, bijvoorbeeld door aantasting van de eindtoets en het centraal examen, en overproductie van diensten en netwerken gericht op innovatie, scholing en begeleiding.

Drie vormen van sturing door de overheid

Het thema hefboomen vestigt verder de aandacht op het besturend vermogen van de overheid bij het uitvoeren van maatregelen om de effectiviteit van het onderwijs op peil te houden of te verbeteren. In de besturingstheorie wordt het gamma aan bestuurlijke maatregelen samengevat als doelsturing, adaptieve (structuur doorbrekende) sturing en routine sturing.

Wat betreft *doelsturing* is in deze studie benadrukt dat het accent ligt op het vakgerichte kerncurriculum. De research evidentie die is samengevat is hier ook op gebaseerd. Ik ga ervan uit dat vakgericht onderwijs de kern is van het funderend onderwijs, maar in het rapport worden impulsen genoemd die dit in twijfel trekken. Als het erom gaat een antwoord te vinden op tegenvallende onderwijsprestaties in basisvakken is duidelijkheid van prioriteiten een eerste vereiste. Misschien is er een nationaal debat over deze prioriteiten nodig. Hoe groot is eigenlijk de politieke wil om de “basiskwaliteit” te verbeteren? Als de prioriteit van het vakgerichte onderwijs wordt herbevestigd, is de boodschap van dit rapport om te zorgen dat er voldoende tijd voor wordt ingeruimd en dat er voldoende leraren zijn die goed zijn toegerust om het vakgerichte onderwijs op een effectieve wijze te geven. In het rapport is deze boodschap verwoord als een eerste scenario tot verbetering gericht op arbeidsvoorwaarden, rekrutering en opleiding van leraren. *Adaptieve sturing* vestigt de aandacht op de structuur van het onderwijsstelsel. In het Nederlandse stelsel is de autonomie van scholen een historisch gegeven. Daarnaast heeft de overheid een aantal grondwettelijke bevoegdheden die kunnen zorgen voor checks & balances, onder meer door het bestaan van centrale examens en de onderwijsinspectie. Internationaal wordt Nederland geprezen om deze combinatie van autonomie en (output) standaardisatie (vgl. OECD, 2016). Wel zijn er aanleidingen om het functioneren van het stelsel aan een nader onderzoek te onderwerpen. Leidt de bestaande vorm van delegatie naar besturen en bestuursorganisaties wel tot autonomie aan de basis, de professionele autonomie van leraren?

Wordt output monitoring niet bedreigd door ontwikkelingen als het vermarkten van de eindtoets basisonderwijs en het streven naar een “examen op maat”? Tenslotte zijn er in dit rapport hypothesen

gegenereerd over overproductie van diensten en het verheffen van middelen tot doelen door het georganiseerde “middenveld” van het stelsel. Het ter discussie stellen van deze onderdelen van het onderwijsstelsel zou een onderzoek vergen van de reikwijdte en statuur van een tweede Commissie Dijsselbloem. Deze gedachte is in het rapport nader onderbouwd onder de noemer van het scenario “Dijsselbloem 2”.

Routinesturing heeft betrekking op maatregelen die plaatsvinden bij gegeven doelen en zonder de structuur van het bestuurd systeem te hervormen. Een rode draad in deze studie is het belang van evaluatie en monitoring, enerzijds ter bevordering van effectieve praktijken en anderzijds ter voorkomen van ineffectiviteit en *inefficiency*. Het in stand houden en optimaliseren van “high stakes” toetsingsmomenten in het funderend onderwijs, in relatie tot “curriculum alignment”, is te zien als een spaarzaam maar effectief verbeteringsscenario, in het rapport beschreven als het scenario van Examenvoorbereiding en uitlijning van het curriculum.

Als het erom gaat “wat te doen om effectiviteit te bevorderen” zijn deze scenario’s bescheiden en kapitaliseren ze op bestaande bevoegdheden en middelen van de overheid om te sturen. Wat betreft de ingesloten boodschappen over “wat te doen om ineffectiviteit en inefficiency tegen te gaan” zijn ze ingrijpend omdat ze de bestaande verhoudingen ten opzichte van besturen en bestuursorganisaties ter discussie stellen. Voor beide is de overheid “aan zet”.

1 INLEIDING

1.1 Vragen van de opdrachtgever

De vraagstelling van de opdrachtgever is hieronder weergegeven op basis van citaten uit de offerte-aanvraag.

Overzicht effectiviteitsonderzoek

“Gezien de recente discussies in Nederland over de doelmatigheid en toereikendheid van de bekostiging in het funderend onderwijs, is het debat over onderwijsverbetering in Nederland gediend met een breed actueel overzicht van bevindingen uit internationaal (en nationaal) empirisch onderwijs effectiviteitsonderzoek, toegeschreven naar de Nederlandse situatie”.

Verklaringen voor verschillen in onderwijsprestaties

“De vraag die in de (wetenschappelijke) studie centraal moet staan, is in welke mate verschillen in prestaties tussen landen, scholen, leerkrachten en leerlingen zijn toe te schrijven aan afzonderlijke door het beleid en de scholen te beïnvloeden factoren. Hierbij moet gebruik gemaakt worden van de (internationale) kennisbasis van het bestaande onderwijs effectiviteitsonderzoek. Dit moet zowel meer helderheid geven over het concept onderwijseffectiviteit als zodanig, als de beschikbare kennis hierover breder ontsluiten en op haar waarde schatten”.

Hefbomen op stelselniveau

“Tevens moet worden gekeken naar de invloed van internationale peilingsonderzoeken (TIMSS, PISA, PIRLS). Specifieke aandacht moet daarbij worden gegeven aan factoren die door de nationale overheid te beïnvloeden zijn -de hefboomen op stelselniveau (denk aan sturen op geld of regels) -, waarbij de grote mate van schoolautonomie in Nederland als een belangrijk gegeven moet worden meegenomen. Dat laatste heeft namelijk gevolgen voor verantwoordingsystemen en het toezicht. In de gevraagde studie moet op deze laatste vraag tevens worden ingegaan door verband te leggen met sturings- en kwaliteitsmodellen en effectiviteitsindicatoren op stelselniveau”.

Beleidsscenario's

“De studie moet uitmonden in een waarderende vergelijking van alternatieve beleid scenario's ten behoeve van het sturen op kwaliteit op nationaal niveau, gestoeld op een wetenschappelijke waardering van internationaal (en nationaal) empirisch onderwijs effectiviteitsonderzoek. Tevens moet de studie een breed actueel overzicht van bevindingen uit dit onderzoek bieden, als aanzet tot een mogelijk structureel te onderhouden kennisbasis”.

1.2 Het perspectief om deze vragen te beantwoorden

Het effectiviteitsbegrip

Het perspectief om deze vragen te beantwoorden is onderwijseffectiviteit.

Het effectiviteitsbegrip is gebaseerd op de aanname dat organisaties doelgericht zijn. Volgens de economische visie op het functioneren van organisaties hebben doelstellingen betrekking op de opbrengsten van het primaire productieproces. Doelbereiking wordt meetbaar gemaakt door doelstellingen te formuleren als “verhoopte opbrengsten”. In het onderwijs gaat het daarbij om onderwijsopbrengsten, in de zin van door middel van toetsen of examens te meten leerprestaties, bereikte onderwijsposities (in de zin van een diploma op een bepaald niveau) en uiteindelijk het ‘civiel effect’ op maatschappelijke posities.

Effectiviteit, of doeltreffendheid, wordt omschreven als de mate waarin de gewenste opbrengst wordt bereikt ten gevolge van geleverde input en beïnvloedbare processen. Doelmatigheid of efficiëntie is vervolgens te definiëren als de maximale output tegen de laagst mogelijke kosten. Met andere woorden, efficiëntie of doelmatigheid is effectiviteit met de aanvullende eis dat deze op zo goedkoop mogelijke manier bereikt wordt.

In het effectiviteitsbegrip staat de middel-doelrelatie tussen allerlei maatregelen en methoden om tot optimale doelbereiking te komen centraal. De term “hefbomen” in de titel van deze studie heeft betrekking op door actoren in het onderwijsveld beïnvloedbare middelen en methoden om optimale onderwijsopbrengsten te bereiken.

Een systemisch perspectief

In het overkoepelend conceptuele kader van deze studie (hoofdstuk 2) wordt effectiviteit gezien als een van de meetbare facetten van onderwijskwaliteit. Hierbij worden doelstellingen niet ter discussie gesteld en ligt het accent op methoden en strategieën om gegeven doelen te bereiken. Een van de andere aspecten van het overkoepelende model refereert aan het zodanig afstemmen en zo nodig bijstellen van doelen als gevolg van veranderde omstandigheden in de ruimere maatschappelijke, sociale en culturele context. Dit kwaliteitsfacet wordt aangeduid met de term “responsiviteit”. Het komt in het rapport aan de orde voor zover de focus op basiskwaliteit die gekozen is, ter discussie kan worden gesteld vanuit andere visies op de gerichtheid van onderwijs, wanneer er bijvoorbeeld een pleidooi wordt gehouden voor “brede vorming”.

In de verdere uitwerking van het effectiviteitsbegrip als centraal punt in het conceptuele kader, (Hoofdstuk 2), wordt aangegeven dat de effectiviteitsvraag op verschillende niveaus in het onderwijs gesteld wordt. Leerlingen die effectieve leerstrategieën toepassen staan centraal. Verder zijn de voornaamste actoren leerkrachten op het niveau van de klas, schoolleiders en bestuurders op het niveau van de school als organisatie en boven-schoolse beleidsvoerders op het niveau van de overheid en nationale onderwijsorganisaties. Omdat in deze studie hefbomen op het niveau van de nationale overheid centraal staan, gaat het uiteindelijk om maatregelen en strategieën die op dit niveau kunnen worden genomen. Echter, de gedachte van een systemisch perspectief brengt tot uitdrukking dat de verschillende hiërarchische niveaus met elkaar in verband worden gezien. Dat betekent onder meer dat centrale besturing “via de band” van lagere niveaus kan plaatsvinden en dat er sprake kan zijn van “meta-besturing”, waarbij de overheid toeziet op de beleidsvoering van lagere overheden en scholen. Een ander facet is het concept van “alignment” oftewel het in lijn brengen van verschillende functiegebieden, zoals de afstemming tussen en examens en curricula, en tussen plannen en de implementatie daarvan. Het systemisch perspectief dat in deze studie gebruikt wordt betekent dat we, als het gaat om het bespreken van empirische evidentie over “wat er werkt”, niet alleen kijken naar

systemeffectiviteit, maar ook naar school- en instructie-effectiviteit. “Last but not least” impliceert het systemisch perspectief op onderwijseffectiviteit de visie dat alle hefboomen, methoden, maatregelen en strategieën uiteindelijk moeten bijdragen aan de optimalisering van onderwijsopbrengsten.

De kennisbasis van onderwijseffectiviteit

Ons uitgangspunt is dat empirische evidentie over effectief onderwijs en effectief onderwijsbeleid bepalend is voor verwachtingen over werkzame “hefboomen”. In deze studie baseren wij ons wat dit betreft op meta-analyses, review studies en secundaire analyses van internationale databestanden. Een tweede perspectief betreft de theoretische fundering van de resultaten, als een belangrijk gezichtspunt om tot conclusies en aanbeveling te komen. Bij het analyseren van de empirische kennis behoort ook het maken van een afweging over de mate waarin de resultaten daadwerkelijk onderbouwd zijn, en moet, waar dat nodig is, ook gewezen worden op de beperktheid van de evidentie. Een derde invalhoek om met de empirische evidentie om te gaan is het gezichtspunt van “ineffectiviteit”. De kennisbasis levert zowel informatie over wat er wel werkt als over wat er niet werkt. Inzicht in ineffectiviteit is voor het beleid minstens zo belangrijk als inzicht in effectiviteit.

En, om mistanden te voorkomen, hoewel er in het rapport voorbeelden over het Nederlandse onderwijs worden aangehaald bij de bespreking van meer of minder effectieve ideeën en praktijken, biedt deze studie geen evaluatie van het Nederlandse onderwijs.

Nadere toelichting op het onderscheid “effectief -ineffectief”

De keuze om de onderzoeksresultaten op deze wijze te presenteren gebeurt door mij met enige pijn in het hart. Een meer diepgaande behandeling van de resultaten van onderwijs-effectiviteitsonderzoek laat veel meer nuances, voorwaardelijkheden en onzekerheden zien. Dat is te lezen in een stuk of twintig boeken en vele artikelen die ik over dit onderwerp geschreven heb. Echter, net zoals men er soms voor kiest om in onderzoek extremen met elkaar te vergelijken, in een eerste exploratieve fase om na te gaan welke hypothesen voor verder onderzoek kansrijk zijn, heeft een dergelijke benadering in het kader van deze studie “formatieve” en “communicatieve” voordelen. Formatief in de zin dat de vergelijking effectief/ineffectief niet als het eind maar als het begin van verder denken/onderzoeken/ontwikkelen moet worden gezien. Communicatief in de zin dat het de leesbaarheid en helderheid van de boodschap hopelijk vergroot. Auteurs als Peal (2014) en Guthry (2021) vatten effectieve resp. ineffectieve benaderingen in het onderwijs samen onder de labels “traditioneel” (effectief) en “progressief”. Op deze bijdragen wordt verder ingegaan in hoofdstuk 6 van dit rapport. Zie ook de bespreking van Peal’s boek “Progressively worse” door Paul Kirschner in *Didactief Online* (2020).

Implementatie en besturend vermogen van de overheid

Het overzicht van de resultaten van systeem-, school- en instructie-effectiviteit kan in principe worden gezien als een blauwdruk voor effectief onderwijsbeleid. Maar om deze daadwerkelijk te kunnen realiseren is bestuurlijk vermogen van de overheid gevraagd om beleidsimplementatie mogelijk te maken. De complexiteit van het te bestuderen systeem, in dit geval het funderend onderwijs in Nederland, biedt specifieke mogelijkheden en beperkingen. De thematiek van kennisproductie en kennisverspreiding krijgt momenteel veel aandacht en wordt onder meer voorgesteld als co-constructie van verschillende categorieën betrokkenen. Hier is het perspectief anders en wordt uitgegaan van externe kennisontwikkeling op basis van wetenschappelijk onderzoek en de benutting hiervan door de overheid. De implementatieproblematiek moet niet onderschat worden, maar valt grotendeels buiten het bestek van de deze studie; zie verder hoofdstuk 6.

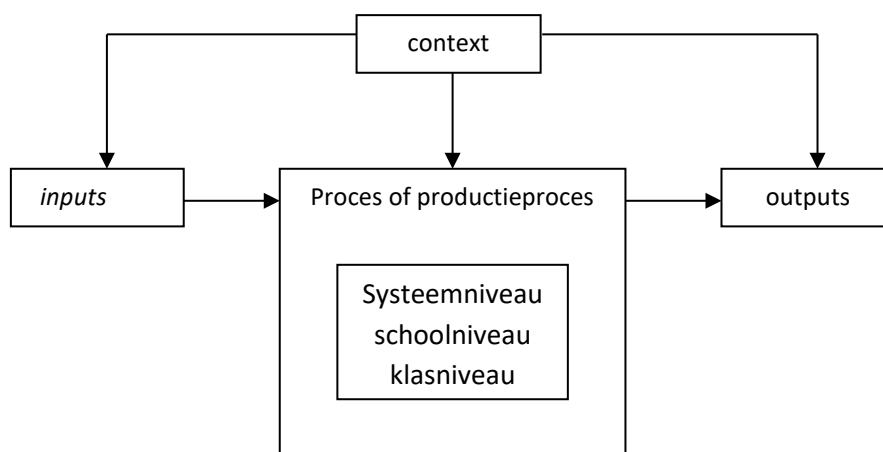
Beleidsscenario's

Er zijn meerdere invalhoeken om op basis van overheidsbeleid en maatregelen van de overheid te proberen het onderwijs op peil te houden of te verbeteren: de initiële opleiding van leraren, de werving en rekrutering van onderwijspersoneel, arbeidsvoorwaarden en rechtspositie, het curriculum voor het funderend onderwijs, de organisatie van het schoolsysteem (onder meer de stratificatie van schooltypes in het voortgezet onderwijs), examens en toetsing, onderwijsinspectie, "governance" en bestuur, financiering, professionele ontwikkeling en ondersteuning van scholen). Kennis van onderwijseffectiviteit kan op al deze terreinen worden gebruikt. We werken dit uit in een beperkt aantal beleidsscenario's en bezien de mogelijkheden tot implementatie van elk van de scenario's in de Nederlandse situatie.

2 EEN OVERKOEPELEND CONCEPTUEEL KADER

2.1 Het basis systeemmodel om onderwijskwaliteit te typeren

Het basisontwerp van onderzoek naar de effectiviteit van onderwijs is het verband tussen hypothetische effectiviteit verhogende maatregelen en de leerprestaties van leerlingen. Het zogenoemde CIPO (Context-Input-Process- Outcome) model, ontleend aan de systeemtheorie, verduidelijkt dit ontwerp (Fig.1). De belangrijkste taak van onderzoek naar de effectiviteit van onderwijs is het aantonen van de invloed van relevante input-kenmerken op output en het openbreken van de “black box”, zodat we kunnen zien welke processen of productiefactoren een doeltreffende werking hebben, naast de invloed van contextuele omstandigheden. Het model van figuur 2.1 kan op verschillende aggregatieniveaus gebruikt worden. In figuur 2.1 wordt dit aangegeven door de drie niveaus in de centrale “black box” van het model: het niveau van een nationaal onderwijssysteem, het schoolniveau en het niveau van de instructie-omgeving, dat vaak aangeduid wordt met klasniveau. De drie niveaus zijn “genest”, dat wil zeggen: de scholen functioneren binnen een onderwijssysteem op nationaal niveau en klassen functioneren binnen scholen.



Figuur 2.1: Een basissysteemmodel voor het functioneren van het onderwijs

Het idee van een onderwijssysteem als een aantal “geneste” lagen, wordt duidelijker wanneer we naar variaties van het basismodel kijken (als weergegeven in figuur 2.1).

Het model dat bij het INES-project van de OECD werd gebruikt, betreft een dergelijk systeem (zie tabel 2.1). Het model heeft twee dimensies: aggregatieniveau van de belangrijkste actoren in het onderwijssysteem en aspecten van de productiviteit van onderwijs, volgens het CIPO- model. (Qua aggregatieniveau worden onderscheiden het nationaal niveau, schoolniveau, het niveau van de klas of leer groep en het niveau van de leerlingen oftewel individuele deelnemers aan het onderwijs. De aspecten van de productiviteit van onderwijs, volgens het CIPO-model zijn resultaten, beïnvloedbare processen en inputs en contextuele condities. Het voordeel van dit model is dat het erkent dat input, processen en output-indicatoren op verschillende niveaus kunnen worden geformuleerd, waaronder het niveau van de individuele leerling.

Tabel 2.1: Het INES-model

	Outputs	Beïnvloedbare inputs en processen	Voorgaande en contextuele omstandigheden
Individuele leerlingen	Bijv. leesprestaties	bijv. effectieve leertijd tijdens leeslessen	bijv. culturele achtergrond
Klas of leergroep	bijv. gemiddelde leesprestaties per klas of per leerjaar	bijv. gestructureerd onderwijs	bijv. klassamenstelling w.b.t. SES
Onderwijsinstellingen	bijv. gemiddelde leesprestaties per school	bijv. onderwijskundig leiderschap	bijv. schoolsamenstelling w.b.t. SES
Het nationale onderwijsstelsel	bijv. gemiddelde leesprestaties op nationaal niveau	bijv. evaluerend vermogen op nationaal niveau	bijv. culturele homogeniteit

Met behulp van de structuur van dit model kunnen de belangrijkste meetbare facetten van onderwijskwaliteit worden weergegeven:

- *Productiviteit van onderwijs* wordt aangegeven door te kijken naar uitkomstvariabelen op verschillende aggregatieniveaus. De welbekende vergelijkingen tussen de gemiddelde prestatieniveaus van verschillende landen in evaluatieve onderzoeken naar onderwijs, zoals TIMSS en PISA, zijn voorbeelden van vergelijkingen van de productiviteit van onderwijs tussen nationale systemen.
- *Effectiviteit van onderwijs*: geeft het netto-effect aan van beïnvloedbare onderwijscondities op uitkomsten, waarbij er gecontroleerd wordt voor achtergrond kenmerken van individuele leerlingen.
- *Gelijkheid van onderwijs* wordt gemeten door de opbrengsten van verschillende groepen leerlingen te vergelijken en na te gaan in hoeverre opbrengsten bepaald worden door achterstandsindicatoren. Bijv. de leesprestaties van meisjes met een niet-Nederlandse culturele achtergrond, de gemiddelde prestatieniveaus van scholen in landelijke gebieden.
- *Efficiëntie van onderwijs*: hierbij gaat het om effectiviteit tegen de laagst mogelijke kosten;
- *Responsiviteit van het onderwijs*; hierbij staat de vraag centraal of er sprake is van een goede afstemming van het onderwijs op de behoeften van de omgeving, bijvoorbeeld de arbeidsmarkt of meer algemeen maatschappelijke verwachtingen over het onderwijs.

2.2 De hiërarchische structuur

Lerende leerlingen als uitgangspunt

Onderwijssystemen kunnen gezien worden als hiërarchisch opgebouwde structuren, met op elk niveau losjes verbonden kernprocessen. Van onderaf gezien, kan het leerproces van individuele leerlingen hierbij als beginpunt worden genomen.

Onderwijs en lesgeven als stimulans van leren

Het onderwijs, dat op de volgende laag van het hiërarchisch systeem ligt, moet gezien worden als stimulans of “booster” die leerprocessen bij leerlingen bevordert. Het idee van een “booster” drukt uit dat de beïnvloeding door een hoger niveau uit het hiërarchische systeem moet worden gezien als een extra impuls om een al goed lopende motor nog iets beter te doen presteren. Uiteindelijk kunnen leerprocessen zelfregulerend zijn, maar in het onderwijs staat structurering ervan centraal. Verschillende gradaties van structurering kunnen op een continuüm geplaatst worden dat reikt van structuur naar onafhankelijkheid. De leeromgeving op het niveau van de klas of leergroepen wordt opgevat als bepaald door directe ingrepen van de docenten en beïnvloed door de leeromgeving of ‘ecologie’ van de klas.

Ondersteunende schoolorganisatie

In meer traditionele visies op de school als organisatie heeft de professionele autonomie van de docent een prominente plaats. Maar ook in recentere visies, die de nadruk leggen op gedeeld leiderschap op school en ‘docentleiderschap’, speelt deze relatieve autonomie nog steeds een belangrijke rol. Scholen worden echter tevens gezien als bestuurde organisaties, waarin een centrale missie, coördinatie, schoolevaluatie en controle, evenals richtlijnen voor het curriculum, het werk van de docenten in banen leiden. Evenals bij de relatie docent-leerling, zien we dus dat de werkrelatie tussen schoolleiders en docenten dezelfde dynamische spanning vertoont tussen richtlijnen en vastgestelde structuren aan de ene kant en zelfstandig functioneren aan de andere kant. Ook op schoolniveau worden klimaat- en ecologische factoren onderscheiden, als meer interactieve dimensies van de schoolomgeving.

Sturing van het onderwijsstelsel door de overheid

De bovenste laag van het systeem wordt gevormd door het nationale of regionale/staats- onderwijsbeleid. De laatste decennia hebben structurele hervormingsmaatregelen met betrekking tot decentralisatie en verantwoordelijkheid in veel landen een belangrijke rol gespeeld. Ook hier worden we weer geconfronteerd met gemengde denkbeelden over controle van bovenaf versus autonomie voor de lager liggende lagen van het systeem. Dit wordt goed weerspiegeld door het concept van functionele decentralisatie, waarbij erkend wordt dat systemen op een bepaald functioneel gebied kunnen centraliseren, bijvoorbeeld met betrekking tot het curriculum, en decentraliseren op andere functionele gebieden, zoals het financiële beheer van scholen.

Relaties binnen het onderwijs als hiërarchisch systeem

Bij de voorstelling van onderwijssystemen als hiërarchische, losjes met elkaar verbonden systemen, kunnen de volgende onderlinge relaties onderscheiden worden:

- Onder de vooronderstelling dat onderwijs op een abstracte manier omschreven kan worden als een productieproces, dat uiteindelijk leidt tot leerresultaten van leerlingen, kan het functioneren op elk niveau gezien worden als een geheel van bestuurd of beïnvloedbare processen en inputs, die samen met de invloed van randvoorwaarden leiden tot outputs.
- Relaties tussen de niveaus worden gezien als directe controlemaatregelen die uitgeoefend worden door een hoger niveau op het onderliggende niveau en als het vormen van al bestaande condities in specifieke 'ecologieën' (dit concept zal later verder toegelicht worden). In het ideale geval wordt het beleid op de hogere niveaus door een feedbackproces op de hoogte gesteld van de resultaten op het lagere niveau.
- Het idee van de losjes verbonden lagen is gebaseerd op de aanname dat de actoren van elk niveau een zekere, of zelfs ruime, autonomie hebben, waardoor het functioneren op een bepaald niveau niet volledig gecontroleerd en vastgelegd wordt door een hoger niveau, maar ruimte laat voor zelfontwikkeling en zelforganisatie.
- Een ideaaltypische interpretatie van een dergelijke losse verbondenheid is het subsidiariteitsprincipe, dat ervan uitgaat dat alles wat redelijkerwijs door een lager niveau uitgevoerd kan worden, niet overgenomen moet worden door een hoger niveau. Opvattingen over onderwijsbeleid, schoolmanagement, onderwijs en leren verschillen hierbij echter wat betreft de definitie van 'redelijkerwijs' en we dienen de analytische en empirische literatuur te raadplegen om meer informatie te krijgen over waar structurering en waar onafhankelijkheid het meest effectief zullen zijn.

2.3 Alternatieve voorstelling en uitwerking van het basismodel

Het basismodel kan verder worden uitgewerkt door de globale structuur te bestuderen, de meest relevante uitkomsten van empirisch onderzoek te bespreken en tenslotte door theoretische opvattingen en verklarende mechanismes te analyseren.

De discussie over de dynamische aspecten van het basismodel, zoals getoond in figuur 2.1, wordt vergemakkelijkt door de invloeden tussen de verschillende niveaus in kaart te brengen. Dit is gedaan in figuur 2.2. De invloeden tussen de niveaus worden aangeduid door middel van gestippelde pijlen die van de hogere niveaus naar de lagere niveaus wijzen. Deze relaties tussen de niveaus kunnen gezien worden als controle, steun en bescherming van een hoger niveau gericht op de kernprocessen van het daaronder liggende niveau.

Opbouw van het model

De figuur volgt de structuur van het CIPO-model, op elk niveau. Naast opbrengsten of outputs worden er verschillende categorieën van invloeden onderscheiden: vooraf "gegeven" condities, aangeduid als "antecedenten", ecologische condities en interventies van de actoren op de verschillende niveaus.

Vooraf gegeven condities (antecedenten)

Deze condities vallen buiten de directe invloedssfeer van de centrale actoren, maar hebben wel invloed op de opbrengsten en de interventies. Op de niveaus van de school, de klas of leergroep en de individuele leerlingen wordt onderscheid gemaakt tussen gegeven condities die bestaan uit invloeden

van een hoger bestuursniveau (antecedenten 1) en kenmerken van de omgeving op hetzelfde niveau en van de kernactoren (antecedenten 2).

Ecologische condities

Een tweede kernelement in figuur 2.2 is de weergave van ecologie als een afzonderlijke klasse van omstandigheden die onderwijsprestaties beïnvloeden. Dit wordt gedaan door een meer expliciete plaats toe te kennen aan deels controleerbare compositie-effecten en hun interactie met algemene omstandigheden die gezien worden als beïnvloedbare variabelen, zoals het schoolklimaat. De erkenning van deze soort contextuele omstandigheden benadrukt de relativering van directe controle in het onderwijs, maar richt de aandacht tegelijkertijd op een categorie van controlemaatregelen van een andere aard, namelijk die van selectie, toelating, groeperen en het kiezen van bepaalde docenten voor bepaalde subgroepen leerlingen en op culturele aspecten die met de samenstelling van groepen leerlingen en docenten samenhangen. De kracht van de interactie tussen compositie en beïnvloedbare variabelen – zoals bij de gezamenlijke impact van een ‘goed’ relationeel klimaat op school en een gemiddeld hoog SES van de leerlingen op de school of in de klas op de prestaties van de leerlingen – is ook relevant voor onderwijsgelijkheid. Interessant is – zoals een secundaire analyse van PISA 2000 aantoonde – dat in een aantal landen de correlatie tussen de gemiddelde SES en ‘goede’ onderwijsomstandigheden veel sterker is dan in andere landen. Dit duidt erop dat deze laatste landen minder goed zijn in het creëren van gelijkheid, in de zin van onafhankelijkheid van sociale herkomst (Luyten, Scheerens, Visscher et al., 2005).

Actor gestuurde interventies

Hierbij gaat het om door de kernactoren op elk niveau beïnvloedbare condities, waarvan wordt aangenomen dat ze doeltreffend zijn om opbrengsten te verbeteren. Overheidsbeleid dat op kwaliteitsverbetering is gericht, schoolmanagement en -organisatie om het primaire onderwijsproces te ondersteunen en te beschermen, effectieve instructie door leerkrachten en opbrengstgerichte leerprocessen bij leerlingen.

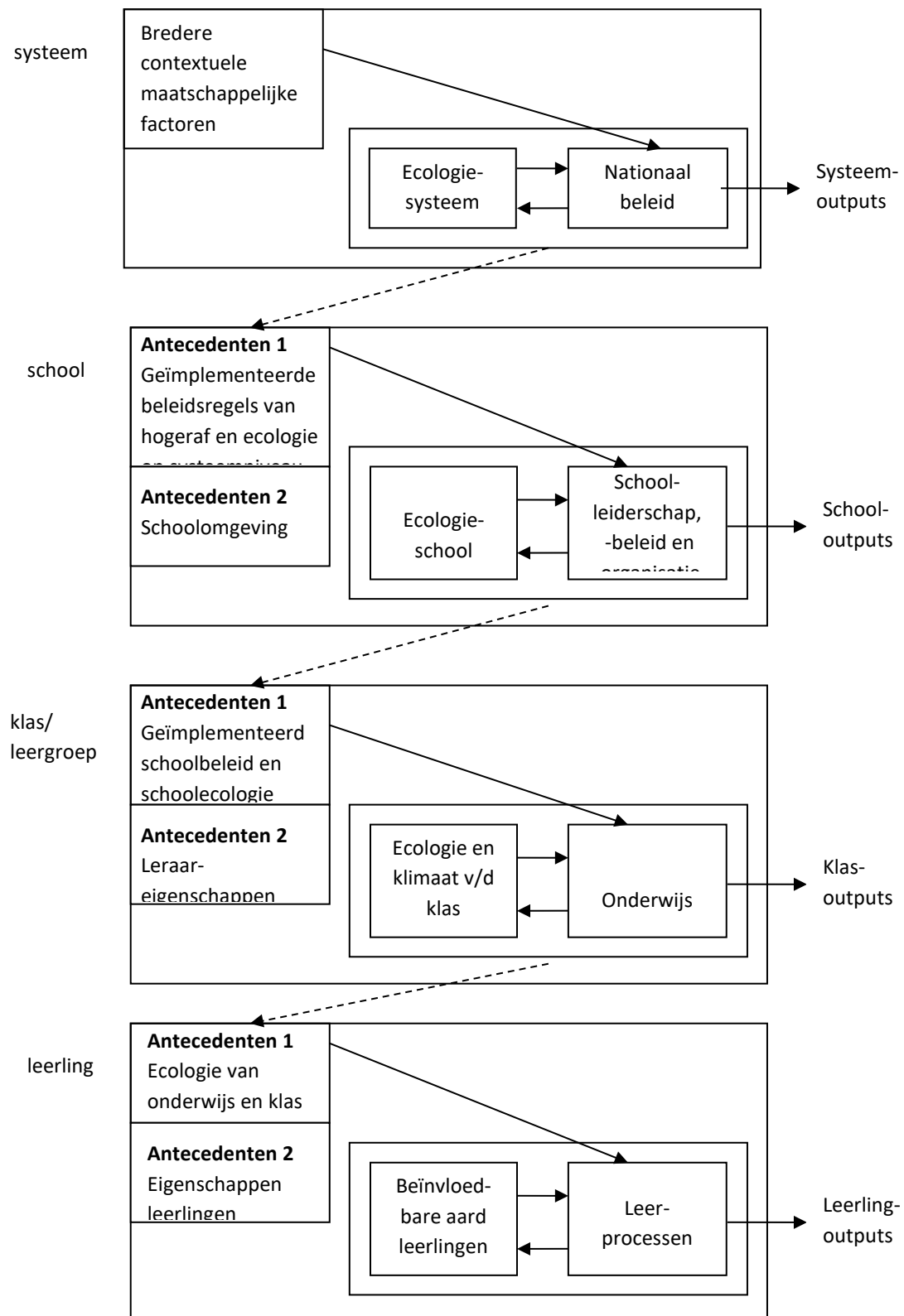
Controle versus losse koppeling

Door het onderwijs op deze manier voor te stellen, wordt uitgedrukt dat de hogere niveaus invloed uitoefenen op de lager liggende niveaus terwijl de hogere niveaus op hun beurt, door de terugkoppeling van feedback, beïnvloed worden door wat er op de lagere niveaus gebeurt. De empirische evidentie die is gevonden in wetenschappelijk onderzoek – en verder beschreven zal worden in de hierop volgende hoofdstukken – onderstreept dat de niveaus relatief onafhankelijk van elkaar zijn. Ondanks het concept van controle door de hogere niveaus hebben de lagere niveaus aanzienlijke beslissingsvrijheid over hun kernprocessen. Zij hebben dus aanzienlijke autonomie en dit wordt weerspiegeld door de losse verbindingen tussen de verschillende hiërarchische niveaus.

De mate van controle door het hoger liggende niveau op het lager liggende niveau is een essentiële kwestie op alle niveaus. Op systeemniveau draait het om een doeltreffend ontwerp van functionele decentralisatie, wat betekent dat - eventueel afhankelijk van de bredere context - een bepaalde mate van centralisatie op een bepaald functioneel gebied (bijv. het curriculum) en decentralisatie op een ander gebied (bijv. financieel beheer) het beste werkt. Op schoolniveau draait het om de mate van deelname aan het besluitvormingsproces, ofwel ‘gedeeld leiderschap’ en op klassikaal niveau draait het om het evenwicht tussen een strak gestructureerde didactische aanpak en meer open onderwijs en leersituaties die zelfregulerend leren moeten aanmoedigen. Structuur versus onafhankelijkheid is de rode draad die door alle beleids- en onderzoek agenda’s op het gebied van onderwijs loopt.

Faciliterende relaties tussen de niveaus

De focus van het model is het stimuleren van leren en leerprestaties. De relaties tussen de verschillende lagen moeten worden gezien als het faciliteren van leren bij leerlingen door didactische handelingen van leerkrachten, het ondersteunen van leerkrachten door de schoolleiding en de schoolorganisatie en het faciliteren en stimuleren van effectief functionerende scholen door de overheid. De aanname hierbij is dat deze focus op leren en leerprestaties leidend is bij keuzes, die uiteraard ook normatief bepaald zijn. Helderheid van standaarden en “targets” is hierbij van groot belang, maar daarnaast is commitment aan prestatiegerichtheid een belangrijk gezichtspunt, dat allerm minst onomstreden is.



Figuur 2.2: Geïntegreerd meerlagig onderwijsmodel.

De gestippelde pijlen van het ene niveau naar het volgende niveau weerspiegelen de invloeden tussen de niveaus; de terugkoppeling van feedback loopt van de uitkomsten van elk niveau naar het vak met ecologie en actief beleid op elk niveau en van lagere naar hogere niveaus. Deze terugkoppeling wordt niet getoond omdat het schema hierdoor onoverzichtelijk zou worden.

Extra besturingsniveaus

Tot dusver is er geen rekening gehouden met besturingsniveaus tussen de centrale overheid en scholen. In sommige landen berust veel uitvoeringsmacht bij regionale overheden of gemeenten. Nederland heeft traditioneel een sterk ontwikkeld middenveld van onderwijsorganisaties, dat grote invloed heeft. Voorheen waren dit de op denominatie berustende onderwijskoepels en pedagogische centra, tegenwoordig wordt het middenveld bepaald door de Sector Raden. Op de voor- en nadelen van dit gegeven in het Nederlandse onderwijs zal verderop in dit rapport nader ingegaan worden.

3 WAT WERKT ER OP HET MICRO-NIVEAU (LEREN EN INSTRUCTIE)

In deze studie wordt getracht een breed overzicht te geven van de kennisbasis op het gebied van onderwijseffectiviteit. Tegelijkertijd is de scope van dit project zeer beperkt. Dat betekent dat er niet de ruimte is om uitgebreide toelichting te geven op specifieke onderdelen. Voor die specificatie wordt verwezen naar literatuur, waar wel uitgebreid is ingegaan op de definities van kernvariabelen, methodologische aspecten van het onderzoek en andere details. In sterke mate wordt hierbij gebruik gemaakt van eerdere publicaties van de auteur. Er wordt gekozen voor de volgende presentatie. Eerst wordt een overzicht gegeven van de belangrijkste variabelen die resp. op micro, meso, en macroniveau zijn onderzocht. Dit gebeurt in overeenstemming met de structuur van het overkoepelende model, weergegeven in Fig. 1.3 " door elk van de niveaus afzonderlijk "in te vullen" met de meest onderzochte variabelen. In dit hoofdstuk behandelen we dus het microniveau, met ingevulde deelmodellen van leren en instructie. Omdat geen ruimte is om iedere afzonderlijke variabele te behandelen, bepreken we vervolgens, in de tweede plaats, de voornaamste categorieën van variabelen. Ieder hoofdstuk wordt afgesloten met een synthese van de empirische ondersteuning van de diverse variabelen-categorieën en hun theoretische onderbouwing.

3.1 Overzicht van variabelen bij leren

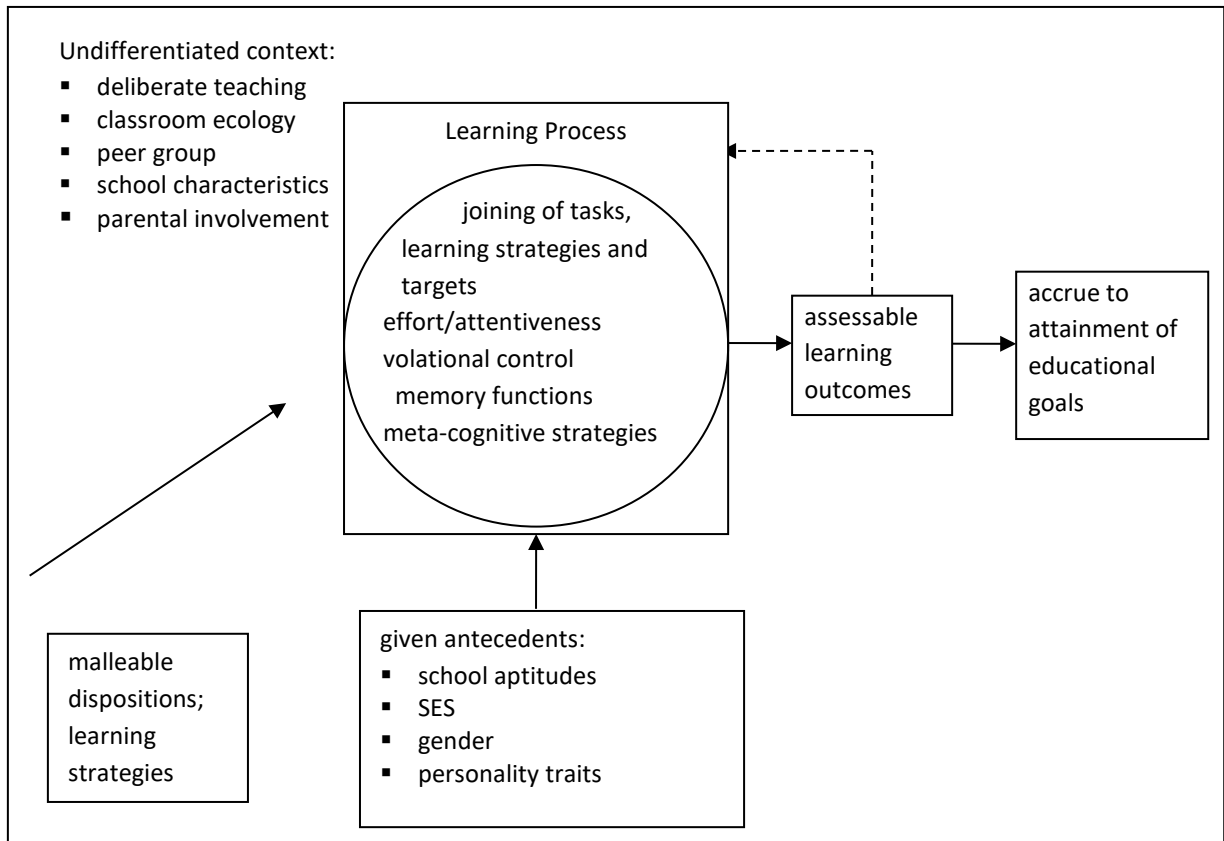
Gezien als "product" kan leren opgevat worden als toegenomen kennis en vaardigheden en eventuele ontwikkeling van waarden en normen. Als we kijken naar het leerproces, dat al dan niet een reactie kan zijn op instructie, zijn de volgende aspecten in het geding:

- Gestimuleerd worden om een leertaak uit te voeren, een probleem op te lossen of een bepaalde houding in te nemen.
- Het beschikken over voldoende energie en motivatie om tot actie over te gaan.
- Gebruik van het kortetermijngeheugen bij confrontatie van nieuwe stimuli.
- Gebruik van langetermijngeheugen bij verwerking van kennis.
- Het inkaderen van de leertaak op basis van eerder opgedane kennis en ervaringen.
- Het uitvoeren van cognitieve operaties op leerinhouden, die kunnen variëren van eenvoudige memorisatie en reproductie tot meer complexe toepassingen, zoals het oplossen van problemen, evalueren en het bedenken van creatieve oplossingen.
- De mogelijke toepassing van meta-cognitieve strategieën, systematische werkwijzen en zelf-evaluatie.
- Het opleveren van een (voorlopig) eindproduct van het leren, dat voldoende tastbaar is om te kunnen dienen als basis voor reflectie en feedback.

Relevante "gegeven" verschillen, op basis van aanleg in milieu, hebben invloed op het gemak, cq de moeite waarmee leerlingen leren. Hierbij gaat het om intelligentie, motivationele disposities, die milieubepaald kunnen zijn, en voorkeuren voor bepaalde leerstrategieën. In een onderzoek voor de OECD onderscheidten Artelt en anderen (2003) drie dominante leerstrategieën: memorisatiestrategieën, "elaboratie" strategieën en controle strategieën. Elaboratie strategieën laten voorkeuren en bekwaamheden zien tot transfer, door wat eerder in een bepaalde context geleerd is toe te passen op nieuw materiaal. Controlestrategieën staan voor een systematische werkwijze, reflectie en zelfcorrectie. De auteurs constateerden dat deze laatste strategieën het meeste

succes hadden. Leerstrategieën worden niet beschouwd als zuiver “gegeven” achtergrondkenmerken, maar eerder als disposities die voor een deel door onderwijs gestimuleerd worden.

In tabel 3.1 worden de voornaamste variabelen bij leren, geordend volgens de opzet van het overkoepelende model uit Hoofdstuk 2, weergegeven.



Figuur 3.1: Model van leren, ontleend aan Scheerens, 2016, p.29

In de onderstaande tabel worden meer specifieke variabelen die een plaats hebben binnen dit model, gerubriceerd als achtergrondvariabelen, gedeeltelijk beïnvloedbare disposities en leerprocessen. Leren onttrekt zich grotendeels aan de waarneming van externe observatoren. Vandaar dat het onderscheid tussen zichtbare tekenen van cognitieve activiteit (*overt*) worden onderscheiden van processen die zich in het hoofd van de leerlingen afspelen en die soms worden opgeroepen door hardop denken uit te lokken (*covert*).

Tabel 3.1. Overzicht van variabelen bij leren

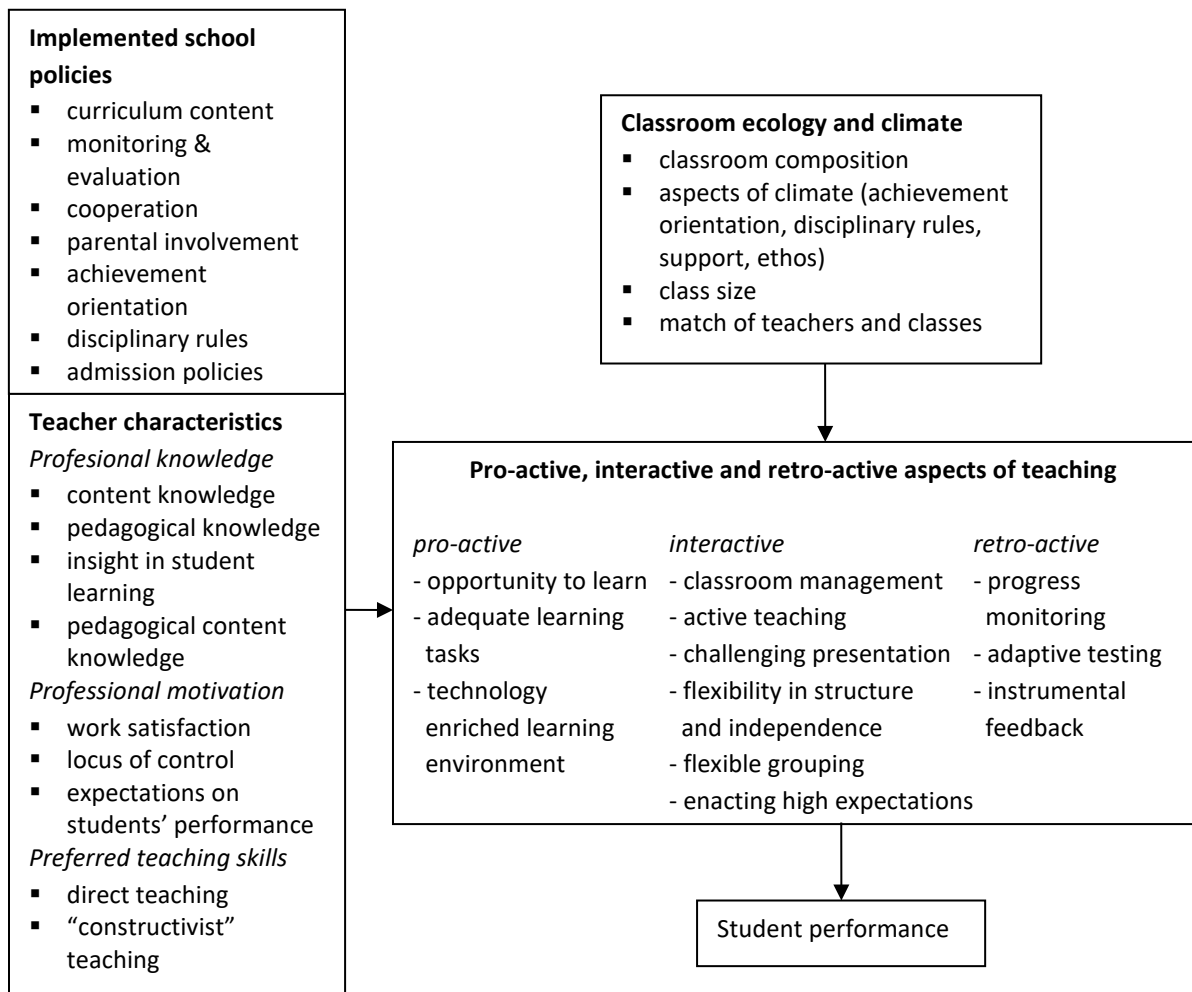
Background variables	Partly teachable dispositions	Learning processes
General intelligence	Learning strategies:	<i>Overt</i>
Scholastic aptitude	▪ memorization strategies	On-task behavior
Socio-economic status	▪ elaboration strategies	Level of engagement
Gender	▪ control strategies	<i>Covert</i>
Immigrant status	▪ domain specific strategies	Self-report on ongoing learning processes (think out loud procedures)
Relevant personality traits, e.g. locus of control,	Subject matter interest	
	Instrumental motivation	
	Persistence	
	Self-efficacy	
	Subject matter based and general academic self-concept	

3.2 Evidentie voor effectiviteit van de voornaamste categorieën van beïnvloedbare variabelen bij leren

In het onderdeel over lesgeven en instructie komen de voornaamste variabelen aan bod, waarmee invloed op het leren wordt uitgeoefend. Zo is een aantal onderwijsstrategieën erop gericht om de leerlingen bij de les te houden en hen voldoende tijd en gelegenheid te bieden om te kunnen leren. Ook de gedeeltelijk beïnvloedbare factoren in de tweede kolom van tabel 1 vragen om interventies van leerkrachten in de sfeer van het stimuleren van motivatie en het bieden van cognitieve en emotionele ondersteuning van leerlingen. Wij zullen deze strategieën bespreken in het volgende onderdeel van dit hoofdstuk dat betrekking heeft op lesgeven en instructie. Ook de evidentie over deze onderwijsingrepen op het terrein van het bieden van gelegenheid tot leren en ondersteuning van leerprocessen komt in het volgende hoofdstuk aan de orde.

3.3 Overzicht van variabelen bij onderwijzen

In figuur 3.2 plaatsen we de variabelen die van invloed zijn bij onderwijzen in de structuur van het overzichtsmodel uit figuur 2.3. Het geïmplementeerde schoolbeleid behoort tot de condities die bepaald worden vanuit een hoger besturingsniveau, in dit geval de school. Leerkrachtkenmerken hebben te maken met gegeven eigenschappen van de kernactoren op een bepaald niveau. Klasse ecologie en klasklimaat verwijst naar gedeeltelijk gegeven en gedeeltelijk beïnvloedbare condities, op het niveau van de klas. Procesmatige onderwijsinterventies zijn gerubriceerd als proactieve, interactieve en retroactieve interventies.



Figuur 3.2: Model van onderwijs, ontleend aan Scheerens, 2016

In de onderstaande tabel 3.2 wordt een deel van de variabelen nog eens samengevat, onderverdeeld als achtergrondkenmerken van leerkrachten, klasse-ecologie en klasklimaat en onderwijsprocessen.

Tabel 3.2 Overzicht van variabelen relevant voor de effectiviteit van onderwijzen, ontleend aan Scheerens, 2016

Teacher background characteristics	Classroom ecology and climate	Teaching processes
<p><i>Professional knowledge</i></p> <ul style="list-style-type: none"> - content knowledge - pedagogical knowledge - insight in student learning - professional content knowledge <p><i>Professional motivation</i></p> <ul style="list-style-type: none"> - work satisfaction - locus of control <p><i>Preferred teaching styles</i></p> <ul style="list-style-type: none"> - direct teaching - “constructivist” teaching 	<ul style="list-style-type: none"> - class size - classroom composition (average and heterogeneity) - match of teachers and classes - aspects of classroom climate, achievement orientation, discipline, support, ethos - teacher expectations on students’ achievement 	<p><i>Pro-active strategies</i></p> <ul style="list-style-type: none"> - opportunity to learn - selection and design of adequate learning tasks - technology enriched learning environments <p><i>Interactive strategies</i></p> <ul style="list-style-type: none"> - classroom management aimed at optimizing active learning time and opportunity to learn - optimizing structure and independence in teaching - learning to use learning strategies - allowing for manageable adaptivity in teaching - active teaching, diversity in preparation formats - a challenging presentation; cognitive activation; - enacting high expectations <p><i>Retroactive strategies</i></p> <ul style="list-style-type: none"> - setting realistic motivating standards - progress monitoring and assessment - adaptive testing - instrumental feedback

In de literatuur bestaan meerdere overzichten van met name de variabelen in de derde kolom van Tabel 3.2, die betrekking hebben op lesgeven en instructie. Deze lijsten komen redelijk overeen met het hier gegeven overzicht. Een voorbeeld daarvan is van Klieme, (2012), dat geïnspireerd is op een bijdrage van Good, Wiley, and Floresz, (2009). Dit overzicht bestaat uit 4 hoofdcategorieën van variabelen: Aanbod van leerinhouden en structuur, klasse-management, een ondersteunend klasse klimaat en cognitieve activering. In vergelijking met het hier gepresenteerde overzicht krijgt “cognitieve activering” een prominenter plaats. Verderop zullen we nadere aandacht besteden aan dit concept.

3.4 Een beperkt aantal theoretisch onderbouwde dimensies van effectief onderwijzen

Er wordt uitgegaan van een vijftal dimensies, die aansluiten bij het overzicht van variabelen in tabel 3.2. Wat de “procesdimensies” van onderwijzen betreft, zijn deze dimensies van een hoger abstractieniveau dan de variabelen die in de vorige paragraaf zijn benoemd en kunnen worden beschouwd als achterliggende of “latente” factoren. We bespreken deze dimensies onder de volgende labels:

Het belang van de leerkracht en “pedagogical content knowledge”

Curriculum alingment

“Structuur en onafhankelijkheid” bij het onderwijzen

Cognitieve en motivationeel/emotionele ondersteuning van leerlingen

Evalueren en feedback

De voordelen van het beschrijven van dimensies op dit abstractieniveau zijn dat de dimensies een synthese bieden van een sterk gedifferentieerde veelheid van empirische onderzoeksuitkomsten en dat ze in meerdere of mindere mate theoretisch onderbouwd zijn. Daarbij komt dat de dimensies als “beleidsrijk” kunnen worden beschouwd en aangrijpingspunten bieden voor beleid dat gericht is op de verbetering van onderwijsprestaties.

Er is afgezien van het in dit hoofdstuk bespreken de mogelijkheden van onderwijstechnologie en artificiële intelligentie (AI) voor instructie-effectiviteit, als eventuele zesde dimensie. Omdat met name de A.I. toepassingen betrekkelijk nieuw zijn, is de kennisbasis op dit terrein nog minder gevestigd dan het geval is voor de vijf eerder besproken dimensies. Dit terrein wordt echter wel gezien als bijzonder relevant voor verder onderzoek; in hoofdstuk 6 (par 6.7) gaan we hier nader op in. In Annex 3 wordt een aanzet gegeven en worden illustratieve onderzoeken vermeld.

3.4.1 *Het belang van de leerkracht en “pedagogical content knowledge” (pck)*

Het lijkt een gemeenplaats om te zeggen dat goede leerkrachten de basis vormen voor goed onderwijs. Het beleid kan hieraan bijdragen door te zorgen voor voldoende ambitieuze beroepsprofielen, goede lerarenopleidingen en aantrekkelijke arbeidsvoorwaarden. Wat onderwijseffectiviteitsonderzoek kan bijdragen is een nadere precisering van “het leerkracht effect” en inzicht in de vraag wat een goede leraar maakt. In dit verband is interessant wat de onderwijseconomen Hanushek en Rivkin hier over opmerken;

“The extensive investigation of the contribution of teachers to student achievement produces two generally accepted results. First, there is substantial variation in teacher quality as measured by the value added to achievement or future academic attainment or earnings. Second, variables often used to determine entry into the profession and salaries, including post-graduate schooling, experience, and licensing examination scores, appear to explain little of the variation in teacher quality so measured, with the exception of early experience. Together these findings underscore explicitly that observed teacher characteristics do not represent teacher quality”. (Hanushek & Rivkin, 2010, p 267). Het is van belang om in te zien dat er in dit citaat sprake is van twee typering van leerkrachteffectiviteit. Als de auteurs het hebben over een aanzienlijke variatie in de kwaliteit van leerkrachten dan bedoelen ze daarmee hoeveel het uitmaakt om door een bepaalde leerkracht onderwezen te worden, zonder dat daarbij gespecificeerd is waarin goede leerkrachten verschillen van minder goede leerkrachten. Het is een “all-in”, “bruto”-effect dat bepaald wordt door de variantie in leerprestaties tussen de leerkrachten die leerlingen les hebben gegeven. De tweede typering van leerkrachtkwaliteit gaat wel in op de vraag welke kenmerken goede van minder goede leerkrachten onderscheiden.

Onderzoek naar leerkrachtheffektiviteit volgens deze tweede typering heeft plaatsgevonden op drie brede terreinen. Economen en sociologen hebben vooral gekeken naar formele kwalificaties en ervaringen evenals arbeidsvoorwaarden (met name salaris). Het effect van deze kenmerken blijkt in geïndustrialiseerde landen bescheiden te zijn, maar is ongeveer twee keer zo groot in ontwikkelingslanden. Psychologen hebben zich verdiept in de invloed van persoonlijkheidskenmerken (bv Darling- Hammond, 1999). Ook deze benadering laat bescheiden of geen effecten zien. In meer recent onderzoek is aandacht besteed aan de kennis van leerkrachten, waarbij onderscheid gemaakt is tussen pedagogische en vakinhoudelijke kennis. Resultaten zijn wisselend, soms heeft pedagogische kennis een groter effect, terwijl andere onderzoeken een groter effect zien van inhoudelijke kennis (vgl. bijvoorbeeld Wayne and Youngs (2003), Bloemeke et al., 2016).

Een belangrijke bijdrage is de introductie van het concept "*pedagogical content knowledge*" (hier verder af te korten als pck) door Lee Shulman (1986). Hij maakt een onderscheid tussen kennis over leerstof en leerstofgebonden didactische kennis. In zijn eigen woorden "*pedagogical content knowledge is knowledge which goes beyond knowledge of subject matter per se to the dimension of subject knowledge for teaching*" (ibid, p 9). Pck heeft betrekking op de keuze van onderwerpen en topics, bruikbare presentatie, analogieën, illustraties, voorbeelden, uitleg en demonstraties. Ook wordt stilgestaan bij de vraag wat het onderwijzen van bepaalde onderwerpen gemakkelijk of moeilijk maakt voor bepaalde leerlingen, als gevolg van bepaalde concepties of misconcepties bij hen. De kern van het concept is dat leerkrachten fundamentele kennis bezitten over de inhoud en structuur van het vak dat ze onderwijzen. De aanvullende kennis en vaardigheden zijn inzicht in wat hun leerlingen al kennen en kunnen en in effectieve strategieën om bepaalde leerstof over te brengen. Pck is een multidimensionaal concept dat niet eenvoudig te operationaliseren en te meten is. Een goed voorbeeld wordt gegeven door Baumert, Blum et al. (2005) voor het wiskundeonderwijs. Zij werken daarbij met een assessment instrument dat bekend staat als "*situational judgment test*". Hierbij worden praktijksituaties uitgebeeld en bijvoorbeeld op basis van video-opnames of verbale typering voorgelegd aan respondenten, die deze beoordelen. Wat betreft de weinige studies die het effect van pck op leerprestaties hebben nagegaan, worden enkele bronnen genoemd in Scheerens, 2016, p. 30. De resultaten zijn positief. Dit blijkt ook uit een geavanceerde studie van Baumert, Kunter et al. (2010).

Het pck concept kan worden gezien als een uitwerking van het model van didactische analyse (De Corte, et. al, 1974, Bloom et al., 1971). Hierbij is eveneens, net als bij pck, sprake van het bij elkaar brengen van vakgebonden leerinhouden, oefenen van psychologische operaties en onderwijsstrategieën. Scheerens, (2016, 2021) spreekt in dit verband over de "*elementaire onderdelen van onderwijzen*". Deze hebben niet alleen betekenis als een model van kennis waar leerkrachten over zouden moeten beschikken, maar vormen ook de basis voor het formuleren van taxonomieën van onderwijsdoelstellingen, het plannen van curricula en lessen, het lesgeven³ zelf, de evaluatie van onderwijs en het assessment van leerprestaties. Deze onderwijskundige basistheorie biedt de mogelijkheid om het onderscheid tussen het leren van specifieke inhoudsgebonden kennis en vaardigheden als een continuüm te zien, in plaats van een categorische tegenstelling. Verdere uitwerking in Scheerens, 2016, pp 30- 32). Het onderscheid in drie elementaire onderdelen van onderwijzen heeft eveneens betekenis voor andere basisdimensies, zoals de typeren van "*curriculum alignment*" en OTL, en structurering van onderwijs (op basis van de inhoudelijke structuur van vakken en het onderscheiden van cognitieve aspiratieniveaus). Waar het concept pck vooral praktische betekenis heeft voor de opleiding van leerkrachten, heeft de basistheorie implicaties voor curriculumontwikkeling, lesplanning en -monitoring en toets-constructie.

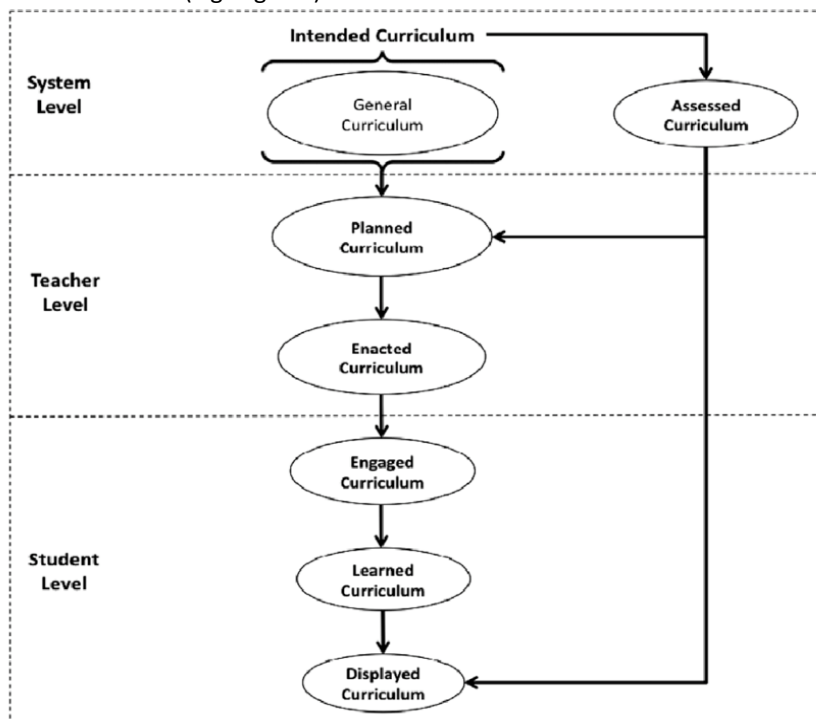
³ Hierbij kan effectief lesgeven worden gezien als "*enacted pck*"

In het onderwijsdebat wordt met het pck-concept het belang van vakgebonden inhoudelijke kennis herbevestigd, als tegenhanger van het constructivisme en de “skills” beweging.

3.4.2 “Gelegenheid om te leren” (OTL) en curriculum-alignment

Begripsbepaling

De eenvoudigste typering van OTL is de aansluiting tussen de aangeboden en de getoetste leerstof. Zeker wanneer er veel afhangt van de uitkomsten van de toets, zoals in het geval van examens, kan het zelfs als een recht van ouders en leerlingen worden beschouwd dat het onderwijs goed voorbereid op de toets. Anderson (2004) definieert OTL dan ook kort en bondig als “the match between what is taught and what is tested”. Een uitbreiding van dit klassieke concept bestaat eruit dat er een tijdsindicator wordt toegevoegd, in de zin van onderwijstijd die aan specifieke leerstofeenheden wordt besteed. Nog verder gaat de uitbreiding van het concept, wanneer er ook een beoordeling van de kwaliteit van het onderwijs als indicator wordt toegevoegd. Voor veel auteurs gaat die uitbreiding te ver, omdat OTL dan in feite samenvalt met onderwijseffectiviteit “als geheel”. Het meeste onderzoek beperkt zich tot de sfeer van het curriculum. De curriculumtheorie onderscheidt het beoogde curriculum, het geïmplementeerde curriculum en het gerealiseerde curriculum. Bij het beoogde curriculum zijn doelstellingen, eindtermen en standaarden de belangrijke ankerpunten. Het geïmplementeerde curriculum is het gegeven onderwijs, dat in lijn met het beoogde curriculum geacht wordt te zijn. Het gerealiseerde curriculum wordt vaak opgevat als de leerresultaten die de leerlingen behalen na het gevolgd hebben van het geïmplementeerde curriculum. Kurz e.a. (2010) breiden de keten nog verder uit, door onder meer een “gepland” curriculum en een door leerlingen ondergaan en geleerd curriculum te onderscheiden (vgl. Fig. 3.3).



Figuur 3.3: Het “Intended Curriculum Model”, naar Kurz e.a., 2011, p 4⁴

⁴ Bron: Handbook of Accessible Achievement Tests for All Students: Bridging the Gaps Between Research, Practice, and Policy (p. 101), by A. Kurz, 2011, New York: Springer

Voor het bezien van de gehele keten, zoals weergegeven in Fig.3.2, wordt de term “curriculum alignment” (letterlijk uitlijning) gebruikt. OTL als een conditie van effectieve instructie richt zich op een bepaalde schakel, namelijk de overeenstemming van het geïmplementeerde curriculum met het beoogde curriculum. In theorie wordt bij de vaststelling van OTL dus een vergelijking gemaakt tussen het gegeven onderwijs en de doelstellingen en standaarden die het beoogde curriculum bepalen. In de praktijk wordt vaker een kunstgreep toegepast, namelijk door het gegeven onderwijs te vergelijken met de inhoud van het finale assessment. Leerlingen of docenten geven dan per toets-item (of categorie van items) aan of de inhoud al dan niet (of in bepaalde mate) in de les is behandeld. Deze kunstgreep berust op de veronderstelling dat het assessment, oftewel de finale toetsing, de standaarden uit het beoogde curriculum dekt. Hieraan wordt tegemoetgekomen wanneer de toetsen inhouds- of doelstellingsvalide zijn, wat uiteraard de bedoeling is. In de theorie over alignment wordt de overeenstemming tussen de standaarden en de eindtoetsen of examens aangeduid met de term *horizontale alignment*. Daarnaast wordt *verticale alignment* gedefinieerd als de aansluiting tussen de curriculum types op de verschillende niveaus. OTL is dus een vorm van verticale alignment, namelijk tussen het beoogde en het geïmplementeerde curriculum.

Empirische resultaten

Een eerste indruk van de grootte van het effect van OTL op leerprestaties is gegeven in overzichtstabel 3.3 verderop. In aanvulling daarop worden nog enkele resultaten vermeld uit studies gebundeld in Scheerens, 2017a. In de eerste plaats werd gekeken naar eerder uitgevoerde meta-analyses. Marzano (2003) vermeldt een gemiddelde effectgrootte, gebaseerd op een aantal oudere meta-analyses, van $d=.88$. Twee, eveneens oudere, meta-analyses hadden betrekking op het effect van toets voorbereiding. Bangert-Drowns et al. (1983) rapporteerden een effectgrootte van $.25$; Messick vond een gemiddelde effectgrootte van $.20$. Een beperkt aantal recentere meta-analyses, van onder meer Kyriakides et al. (2013), Schroeder et al. (2007) en Spada en Tomita (2010) hadden betrekking op “proxys’s” van OTL en lieten gemiddeld een gemiddelde effectgrootte van om en nabij $.30$ zien. Dat is te beschouwen als een betrekkelijk klein effect. In een andere studie in deze bundel (Lamain, Scheerens en Noort, 2017) werd een meta-analyse van het “vote-count” type uitgevoerd. Daarbij gaat de belangstelling vooral uit naar het percentage van positief significante effecten. Dat was in dit geval 44%; qua grootte vergelijkbaar met percentages gevonden voor andere effectiviteit bevorderende factoren, zoals leertijd en prestatiegerichtheid, maar aanmerkelijk groter dan vote-count percentages voor organisatie factoren op schoolniveau, zoals samenwerking tussen leerkrachten en onderwijskundig leiderschap. In een derde studie uit deze bundel werd een secundaire analyse uitgevoerd op het effect van OTL in PISA 2012 en TIMSS, 2011. De effectgrootte bij TIMSS in het vak wiskunde was bijzonder klein ($.074$) terwijl het effect bij PISA 2012, voor het onderdeel formele wiskunde, aanmerkelijk groter was nl $.37$. Al met al is de “evidence base” voor het effect van OTL tamelijk bescheiden. Er zijn betrekkelijk weinig meta-analyses, de onderlinge vergelijkbaarheid van onderzoeken is beperkt, omdat er grote onderlinge verschillen zijn in de operationalisering van OTL en in methoden van data-verzameling. Ook onderzoek dat breder gericht was op verschillende onderdelen van de “alignment keten”, zoals dat van Polikov en Porter (2012) had enigszins tegenvallende resultaten, zeker in vergelijking met de uitkomsten van grote Amerikaanse onderzoeken van William Schmidt en collega’s (Schmidt et al., 2001, Schmidt et al., 2011).

Theoretische onderbouwing

Alignment past in de rationele planningstheorie en de principes daarvan gelden ook voor OTL als onderdeel van systemische alignment. Systematische planning, als pro-actieve structurering, is gebaseerd op het principe “denk voor je doet” of “bezint eer ge begint” om het wat klassieker aan te geven. Het is niet goed denkbaar dat onderwijs gegeven wordt zonder enigerlei toepassing van dit principe. Het gaat hier om de toevoeging dat de planning systematisch verloopt, om de kwaliteit van het onderwijs te verbeteren. Idealiter berust planning om OTL te vergroten het op het op elkaar betrekken van onderwijsdoelen, het onderwijs dat hierop gericht is en het assessment, waarin wordt nagegaan of de doelen in de zin van gerealiseerde leeropbrengsten bij leerlingen bereikt zijn. Het is denkbaar dat de planning geheel berust op de lesvoorbereiding van individuele leerkrachten, al dan niet ingebed in school- of vaksectie brede initiatieven. In de praktijk spelen leergangen en methoden een belangrijke rol. Het is zeer de vraag in welke mate in de onderwijspraktijk keuzes over het lesgeven en het kiezen van assessment-methoden inderdaad worden gedaan op basis van doelstellingenanalyse. Het is denkbaar dat doelstellingen op het niveau van de school of vaksectie worden aangegeven, waarschijnlijker is dat doelstellingen impliciet of expliciet zijn inbegrepen in de leerboeken en methoden.

Bij de definiëring van OTL aan het begin van deze paragraaf werd al duidelijk dat er smalle en bredere invullingen mogelijk zijn. De basisinvulling van OTL is “content covered”; het percentage van de leerinhouden dat deel uitmaakt van het assessment dat daadwerkelijk onderwezen is. Kurz (2011, p 17) bespreekt het opnemen van indicatoren over de cognitieve vereisten (*cognitive demand*), als een van de “verbredingen” van het concept. Deze toevoeging vestigt de aandacht op wat in paragraaf 3.4.1 is aangeduid als de “elementaire onderdelen van onderwijzen”; hierbij bleek dat onderwijsdoelstellingen en toetsingskaders in feite bepaald worden door combinaties van leerstofinhouden en vereiste psychologische operaties. Deze psychologische operaties zijn met name van cognitieve aard, maar kunnen ook voor niet-cognitieve processen worden aangegeven. Het duidelijkst is dit het geval in taxonomieën van onderwijsdoelstellingen (vgl. Bloom, Hastings & Meadows, 1971). Evengoed als leerinhouden van verschillende complexiteit en moeilijkheid zijn, variëren ook de eisen die aan de psychologische operaties worden gesteld (bijvoorbeeld de “cognitive demand”). Naar mijn mening wint het OTL-concept aan relevantie wanneer het gericht wordt op eenheden die de tweezijdige bepaaldheid door inhouden en “cognitive demands” reflecteren. (Nadere verwijzing naar RTTI-dissertatie). Een finesse uit de planningstheorie is dat planning ook “retro-actief” kan worden toegepast. Daarmee wordt bedoeld dat eerst het assessment wordt uitgewerkt, en dat het onderwijs zich daar vervolgens op richt. Meestal wordt een dergelijke aanpak afgekeurd als “teaching to the test”. Daarbij wordt over het hoofd gezien dat toets- en examenvoorbereiding altijd aan de orde is en dat er zowel acceptabele vormen van toetsvoorbereiding voorkomen als ongewenste praktijken (vgl Popham, 2003, Scheerens, 2017a).

Als het gaat om de praktische en beleidsmatige relevantie van OTL en curriculumplanning, geldt dat dit een aanpak is die minder in het centrum van de belangstelling staat en die zelfs omstreden is. De filosofie van autonomie en “eigenaarschap” verzet zich tegen externe curriculum planning en summatieve toetsing. In de internationale literatuur over “schoolverbetering” wordt zwaar gekapitaliseerd op organisatieontwikkeling en leiderschap en wordt de curriculum-kolom stiefmoederlijk behandeld of over het hoofd gezien. Vanuit de kritische pedagogiek worden rationele technieken als planning en evaluatie afgewezen, omdat ze de professionele autonomie van leraren zouden aantasten

(vgl Biesta, 2017). In concepten als dat van “teacher agency” lijkt het wel alsof de individuele leerkracht alle wielen zelf moet uitvinden.

Tegelijkertijd geldt dat de overheid de wettelijke bevoegdheid heeft om eindtermen, standaarden en examens vast te stellen. Dat scheidt de mogelijkheid om, ook in een door autonomie gedomineerd stelsel als het Nederlandse onderwijs, OTL en curriculum alignment te faciliteren. Scheerens en Exalto (2017, 27) signaleren hiervan ook constructieve toepassing en goede praktijken.

3.4.3 **Gestructureerd onderwijs**

3.4.3.1 *“Directe instructie” (d.i.) als kern: didactische ondersteuning*

Het is lastig om een fraaie Nederlandse vertaling van “direct instruction” te geven zonder de term “gestructureerd onderwijs” te gebruiken. In het Engelse gebruik van d.i. is dit een specifiek vorm van gestructureerd onderwijs, naast bijvoorbeeld “mastery learning”, we gebruiken daarom in deze paragraaf “directe instructie” als een specifieke vorm van gestructureerd onderwijs. De term gestructureerd onderwijs is dus breder en omvat ook aan d.i gerelateerde concepten, zoals het “lesgeven-leren-model” van Carroll, 1989) en het al genoemde model van “mastery learning” (Bloom, 1976).

De belangrijkste aspecten van het Carroll model zijn:

- effectieve leertijd die gezien wordt als het resultaat van doorzettingsvermogen en “gelegenheid om te leren” (OTL);
- benodigde netto-leertijd als resultaat van begaafdheid van de leerling, de kwaliteit van het onderwijs en het vermogen van de leerling om het onderwijs te begrijpen.

De verdere didactische uitwerking van deze principes werd door Doyle (1985) als volgt verwoord, als een samenvattende typering van directe instructie:

1. De onderwijsdoelstellingen zijn duidelijk geformuleerd.
2. Het te gebruiken leermateriaal is zorgvuldig gesplitst in leertaken en in een bepaalde volgorde geplaatst.
3. De leraar legt duidelijk uit wat de leerlingen moeten leren.
4. De leraar stelt regelmatig vragen om te beoordelen hoeveel vooruitgang de leerlingen boeken en of ze de stof hebben begrepen.
5. Leerlingen hebben ruimschoots voldoende tijd om te oefenen wat ze geleerd hebben, waarbij veel gebruik wordt gemaakt van ‘geheugensteuntjes’ en feedback.
6. Vaardigheden worden aangeleerd totdat de leerlingen ze automatisch beheersen.
7. De leraar test de leerlingen regelmatig en vraagt de leerlingen verantwoordelijkheid te nemen voor hun werk.

Stockard et al. (2018, 480/481) bespreken een aantal definiërende kenmerken van directe instructie. Zij noemen het principe van “mastery learning” een sleutel element; de verwachting is dat als leerlingen een nieuw begrip volledig beheersen, dit deel gaat uitmaken van bestaande repertoires. Voor alles moet vermeden worden dat concepten op een gebrekkige manier worden aangeboden, want het is gemakkelijker een nieuw begrip te leren dan een verkeerde conceptie af te leren. Verder noemen Stockard et al. een zorgvuldig opgebouwd curriculum een belangrijke basis voor beheersingsleren. Een stapje voor stapje benadering scheidt de mogelijkheid voor continue beloning en bevestiging gedurende het leerproces. Curriculum materiaal moet sterk gestructureerde aanwijzingen bieden aan leraren, voor wat betreft bewoording, opeenvolging, herhaling en toetsing van het aanbod aan de leerlingen. Afzonderlijke inhouden

en vaardigheden worden eerst afzonderlijk behandeld, maar vervolgens gecombineerd in steeds geavanceerdere en complexe toepassingen.

Uit de vroege onderzoeksliteratuur komt onder meer het belang van een ordelijk klimaat in de klas en het vermijden van tijdverlies naar voren (vgl Stalling en Mohlman, 1981). Maar tevens bleek dat gestructureerd onderwijs, volgens het model van directe instructie, niet beperkt is tot het aanleren van eenvoudige feitenkennis, maar ook effectief is bij het aanleren van hogere cognitieve vaardigheden. Zo noemden Collins & Stevens (1982) vijf onderwijsstrategieën die probleemoplossend leren ondersteunen: a) een systematische variatie van voorbeelden; b) tegenvoorbeelden; c) ontlokkingsstrategieën; d) hypothese-identificatiestrategieën; e) hypothese-evaluatiestrategieën. Ook Brophy and Good, 1986, p. 367) beantwoorden de vraag of dit type van zeer gestructureerd onderwijs net zo goed werkt voor het aanleren van gecompliceerde cognitieve processen, in het voortgezet onderwijs positief.

De effectiviteit van directe instructie en gestructureerd onderwijs is breed onderbouwd door empirisch onderzoek. Enkele "highlights" zijn het succes van de op directe instructie gebaseerde variant van het bekende Amerikaanse compensatieprogramma "Follow Through" (Bereiter, 1982), de review van Gage (2008) over "theories on teaching", waarin een vergelijking wordt gemaakt tussen "progressive-discovery-constructivist teaching" en "conventional-direct-recitation teaching", die duidelijk in het voordeel van de laatste benadering uitvalt; en de vergelijking tussen activerend en faciliterend lesgeven van Hattie (2008, 243), met duidelijk sterkere effecten van activerend lesgeven. Ook het grote verschil in effectiviteit tussen "teacher directed science education" in vergelijking met "enquiry-based teaching practices", zoals vastgesteld in PISA 2015 (OECD, 2016, p 65-73), in het voordeel van de eerstgenoemde aanpak, mag in deze niet onvermeld worden gelaten. Last but not least is er de indrukwekkende meta-analyse van Stockard et al. (2015), waarin effectonderzoek naar directe instructie over een periode van 50 jaar een gemiddelde effectgrootte van .50 liet zien.

Theoretische onderbouwing

Kirschner, Seller & Clark (2006) bieden een theoretische verklaring voor de grotere effectiviteit van gestructureerd onderwijs in vergelijking met meer "open" onderwijs. Zij spreken van begeleidend versus minimaal begeleidend onderwijs. Voorbeelden van "minimaal begeleidend onderwijs" zijn benaderingen waarin het leren wordt opgevat als "constructivistisch", "ontdekkend", probleem georiënteerd, ervaringsgericht en "onderzoekend". In hun artikel verwijzen ze naar tal van vergelijkende studies, waarin de effectiviteit van begeleid onderwijs groter bleek dan van vormen van "minimaal begeleid onderwijs". Zij constateren dat de effectiviteit van "minimaal begeleidend onderwijs" geen overtuigende empirische onderbouwing heeft. De verklaring die zij geven is gebaseerd op de cognitieve architectuur van de geheugenwerking. Centraal daarbij is de functie van het korte termijn- of werkgeheugen. Instructie van leerlingen voor wie de leerstof nieuw is moet opslag in het werkgeheugen faciliteren. Bij minimaal begeleid onderwijs, dat bijvoorbeeld gericht is op "ontdekkend leren", is er een groot risico op "cognitieve overladenheid". Hierdoor kan het werkgeheugen niet voldoende functioneren om het uiteindelijke leren, dat berust op het langetermijngeheugen, mogelijk te maken. De remedie van gestructureerd onderwijs is de ondersteuning van de opslag in het werkgeheugen en het vermijden van overladenheid door een aanpak met veel aanwijzingen en oefeningen. Geleidelijk aan kunnen de aanwijzingen verminderd worden door middel van het aanbieden met slechts gedeeltelijk uitgewerkte voorbeelden, om pas in latere fasen te werken in de richting van meer open en complexere opdrachten. Een proces van geleidelijke vermindering van cues dat wordt aangeduid met de term "fading" (vgl. Kirschner, Claessen and Raaijmakers (2018, p 128.)). Eventueel zou deze geleidelijke afzwakking van ondersteuning en begeleiding gezien kunnen worden

als een brug tussen gestructureerd onderwijzen en het bevorderen van ontdekkend leren. In dezelfde geest behandelen Messner & Blum (2019) wat zij aanduiden als “de mythe van open onderwijs”. Zij concluderen dat structurering en “openheid” niet alleen als onverenigbare tegenpolen moeten worden gezien, maar juist als onderling verbonden. Die verbondenheid berust op dezelfde fasering als waarvan sprake is bij het hierboven beschreven “fading”, met dien verstande dat “open” en “onafhankelijk leren” in schoolvakken totaal onhaalbaar wordt geacht zonder daaraan voorafgaande begeleidende instructie door de leerkracht. In de Duitse literatuur over onderwijstheorieën (Baumert et al 2010, Klieme & Rakoczy, 2003, Praetorius et al. 2020)) wordt het concept “cognitieve activering” als een belangrijke dimensie van onderwijzen neergezet. Daarbij wordt gedoeld op aspecten zoals het denken op een hoger niveau, activering van het denkvermogen, diepgaand begrip van de leerstof, betekenisvolle contexten, authentieke instructie, relevantie van de leerstof, toepasselijke taal en taal op een hoog niveau.

Ook variatie in presentatiemethodes wordt genoemd als een manier om cognitieve uitdaging te bieden voor leerlingen. Schreyer, (2020) wijst op de constructivistische achtergrond van dit concept. Ondanks deze achtergrond, dat wil zeggen herkomst uit een theorie, die vanuit het perspectief van directe instructie de plank mislaat, is “cognitieve activering” wel degelijk te verbinden met didactische ondersteuning. Enerzijds vanwege de gemeenschappelijke gebondenheid aan vakgericht onderwijs (in tegenstelling tot algemene “skills”) en anderzijds door dit concept van cognitieve activering nadrukkelijk te situeren in een onderwijsfase, waarin kan worden voortgebouwd op eerdere geleide instructie (vgl. citaat Messner & Blum). In verwijzingen naar de “oudere” literatuur over de effectiviteit van directe instructie werd al vastgesteld dat ook hogere cognitieve vaardigheden op een gestructureerde manier aangeleerd kunnen worden. Dit geldt ook voor het aanleren van meta-cognitieve vaardigheden. Recentere publicaties waarin dit principe wordt bevestigd en uitgewerkt zijn de onderbouwing van “zelf-gereguleerd” leren door de Engelse Education Endowment Foundation (Eef) en publicaties over het onderwijzen van meta-cognitieve vaardigheden van o.m. De Boer, Donker, Costongs, & Van der Werf, (2018) en Donker, (2015).

Voor verdere informatie over de theoretische onderbouwing van “directe instructie” zij verwezen naar Gage, (2008), Van der Werf, (2005), Kirschner et al. (2006) en Scheerens, (2021).

Maatschappelijke relevantie

Als het gaat om de relevantie van de empirische evidentie over “directe instructie” doet zich, volgens auteurs als Stockard et al. (2018) en Peal (2014), een paradoxaal verschijnsel voor: hoewel we redelijk goed weten dat gestructureerd onderwijs beter werkt, in het bijzonder ook voor leerlingen die met een achterstand aan het onderwijs beginnen, wordt telkens weer voorrang gegeven aan varianten van minimaal begeleid en open onderwijs. Deze open benadering heet al 50 jaar “progressief”, terwijl gestructureerd onderwijs als “traditioneel” wordt beschouwd (Vgl. Peal, 2014). Volgens Stockard et al. (2018) moet dit onder meer worden toegeschreven aan de populariteit van het constructivisme, het propageren van “student-led” en “inquiry-based” approaches in curriculum aanbevelingen van de (Amerikaanse) overheid en het vermoeden dat leerkrachten niet van d.i. houden, omdat het hun persoonlijke onderwijsstijl zou belemmeren

3.4.3.2 Ondersteuning in het kader van klassenmanagement, leeromgeving en klimaat

Gestructureerd onderwijs in de betekenis van directe instructie vormt de kern van onderwijzen. Om misverstand te voorkomen, wijs ik er nogmaals op dat deze vorm van onderwijs niet beperkt is tot reproductie van kennis, maar ook geldt voor hogere cognitieve vaardigheden die hierop voortbouwen.

Denk aan wat er gezegd is over het geleidelijk verminderen van de ondersteuning (fading), het gestructureerd werken aan leerstrategieën en “cognitieve activering”. De kritiek op het constructivisme en het ontbreken van positief bewijs van de effectiviteit ervan is gebaseerd op het te vroeg, dat wil zeggen niet gefundeerd in een sterk begeleide onderbouwing, een beroep doen op zelfregulering. Niet op een categorische afwijzing van op constructivisme geïnspireerde werkvormen per se.

In aanvulling op de met directe instructie aangeduide didactische kern van het onderwijs worden in deze paragraaf aanvullende modaliteiten van ondersteuning van leerlingen besproken. Deze modaliteiten zijn eventueel te beschouwen als aspecten van klassenmanagement en klimaat. Wij bespreken deze aanvullende ondersteuningsmodaliteiten onder de volgende noemers:

Klasse-ecologie, het management van leertijd en gelegenheid om te leren, het klasklimaat, cognitieve en emotionele ondersteuning en evaluatie en feedback.

De ecologie van de klas

Onder klasse ecologie worden beschrijvende kenmerken van de klas of leergroep verstaan, zoals grootte en samenstelling evenals de interactie van deze kenmerken met factoren als klimaat en onderwijsstrategie. Andere facetten zijn de matching van leerkrachten met leer groepen, schoolregels en gedragsregels op school. We bespreken hier enkel indicatoren die op dit gebied zijn onderzocht in onderwijs effectiviteitsonderzoek.

De uitvoerige literatuur over de impact van klassengrootte op de leerprestaties wordt niet in detail besproken. De belangrijkste uitkomst is dat reductie van de klassengrootte vrij groot moet zijn voor er enig effect zichtbaar is. Omdat er wel hoge kosten aan verbonden zijn is dit een minder efficiënte strategie om het onderwijs te verbeteren. Decennialange financiële ondersteuning van scholen met hogere percentages “achterstandsleerlingen” in Nederland, waarbij deze middelen vooral zijn gebruikt voor het kleiner maken van de klassen, heeft tot nu toe geen positieve evaluatie-uitkomsten opgeleverd (Driessen, 2018).

De samenstelling van de klas, qua sociaaleconomische achtergrond van de leerlingen, heeft invloed op het klimaat in de klas en vraagt aanpassingen in het lesgeven. Luyten et al. (2005) interpreteerden de mate waarin de schoolcompositie in sociaaleconomische achtergrond de leerprestaties bepaalde als een ongelijkheidsindicator.

In de discussie over een homogene of heterogene klassensamenstelling is het werk van Slavin toonaangevend. Zijn aanbeveling is om leerlingen voor het overgrote deel van de onderwijstijd in heterogene groepen onder te brengen en incidenteel te hergroeperen in homogene niveaugroepen in vakken als lezen en rekenen/wiskunde (Slavin, 1996, p. 164).

In het kader van het bestrijden van onderwijsachterstanden wordt van tijd tot tijd geopperd om een meer strategisch gebruik te maken van het toewijzen van leerkrachten aan leer groepen. De gedachte is dan dat de beste leerkrachten zouden moeten worden toegewezen aan de zwakste leer groepen (Monk, 1989, 1992).

Klassenmanagement

Klassenmanagement werd oorspronkelijk meer specifiek gebruikt voor de maximalisering van effectieve leertijd. Hierbij draait het erom dat onderwijs de leerlingen maximale tijd moet bieden om bezig te zijn met leeractiviteiten, zowel qua instructietijd in het klaslokaal als thuis. De voornaamste aspecten die in de artikelen over de effectiviteit van onderwijs werden genoemd, waren tijd besteed aan de taak, effectief gebruik van instructietijd, leergelegenheid, huiswerk en ‘mastery learning’ (beheersingsleren). Een tweede categorie van instructiekenmerken heeft betrekking op de *organisatie van de klas*. Deze categorie omvat onderwijsaspecten die betrekking hebben op het creëren van een georganiseerde

leersituatie. Aspecten hiervan zijn het toepassen van schoolregels en het zorgen voor een ordelijk verloop van de les en de keuze van werkvormen. Deze laatste categorie kan tevens gericht zijn op *adaptief onderwijs*, het idee dat onderwijs aangepast moet worden aan de kenmerken van de leerlingen. Onder de categorie adaptief leren vallen indicatoren zoals de variatie in het gebruik van onderwijsmethoden, gerichtheid op individuele leerprocessen en het bieden van keuzes. Onderwijsvormen als samenwerkend leren en het incidenteel apart nemen van leerlingen in “een op een”-tutoring-sessies doen eveneens een appel op het klassenmanagement.

Clasklimaat

Clasklimaat wordt omschreven als de algemene sfeer in de klas. Hoofdfacetten van een gunstig, een effectiviteit bevorderend klimaat zijn: goede verhoudingen tussen leerkrachten en leerlingen, waarbij de ondersteunende rol van de leerkracht centraal staat, prestatiegerichtheid, orde en discipline en goede onderlinge verhoudingen tussen de leerlingen. Ondersteuning door de leerkracht is nauw verbonden met structuur en helderheid in het onderwijs en het geven van feedback. Het houden van orde in de klas is een belangrijke voorwaarde om goed les te kunnen geven. Een van de weinige onderwijskenmerken uit internationale studies, die in een relatief groot aantal landen positief samenhangt met prestaties is een ordelijk klimaat. In zijn boek “Progressively worse” schetst Peal (2014) een indringend beeld van de mate waarin gebrek aan orde aan de basis lag van slecht onderwijs in Londense achterstandswijken. Tevens beschrijft hij het succes van een “robuuste” aanpak en strikte disciplinaire regels in dezelfde soort scholen. De “eigenaarschap ideologie” die geassocieerd is met “open”, en in Peal’s typering, “progressief” onderwijs zou een laksere houding ten opzichte van orde en discipline in de hand werken. Ook de nationale cultuur, zoals te typeren volgens de indeling van Hofstede, speelt waarschijnlijk een rol als het gaat om voorkeuren betreffende discipline en “permissiveness”. In de Nederlandse traditie zal men, naar verwachting, minder gauw overgaan tot een strikte aanpak dan in het Verenigd Koninkrijk.

Prestatiegerichtheid wordt mede bepaald door het gebruik van leerdoelen en standaarden, die de basis vormen voor normen bij het assessment. De meer affectieve aspecten van de kwaliteit van de onderlinge relaties wordt medebepaald door de klassensamenstelling en het bieden van emotionele ondersteuning door de leerkracht. Ook prestatiegerichtheid wordt, net als houdingen ten opzichte van discipline en gezag, medebepaald door de nationale cultuur. Een opmerkelijke recente onderzoeksuitkomst, gebaseerd op internationale TIMSS-studie, is dat Nederlandse leerlingen, leerkrachten en schoolleiders het minst prestatiegericht waren van alle deelnemende landen (Meelissen et al., 2020). Ook bleek uit het dissertatie-onderzoek van Neeleman (2019) dat schoolleiders in Nederland meer belangstelling toonden voor pedagogische onderwijsconcepten dan voor kennis over onderwijseffectiviteit.

De meer affectieve aspecten van het clasklimaat worden besproken in de volgende paragraaf.

Cognitieve en emotionele steun bij het lesgeven

Het principe van het bieden van variëteit in aanbiedingsvormen en didactische aanpak is zo oud als er systematisch over onderwijs is nagedacht. *Activerend onderwijs* is een instructiecategorie waarbij het erom gaat leerlingen diverse mogelijkheden te bieden voor actief leren (Slavin, 1995). Deze categorie omvat dan ook aspecten als samenwerkend leren, gesitueerd leren, ‘peer-tutoring’ (waarbij leerlingen - jongere- leerlingen begeleiden), leerlingexperimenten, ‘hands-on’-activiteiten, groepswork, individueel werk, individueel leren en discussies door leerlingen. Activerend onderwijs verschilt niet van gestructureerd onderwijs in de vorm van directe instructie; er wordt echter explicieter op variëteit in aanbiedings- en organisatievormen gefocust, met het oog op het stimuleren van de motivatie en de betrokkenheid van de leerlingen.

Ondersteuning om hogere orde cognitieve operaties bij leerlingen te stimuleren zijn als extensie van directe instructie al ter sprake gebracht in de vorige paragraaf. We resumeren twee benaderingen: het leren van leerstrategieën en “cognitieve activering”, zoals besproken in de Duitse onderwijspsychologie (vgl. Seidel en Steen, 2005).

Leren om leerstrategieën te gebruiken verwijst naar de doelstelling van zelfregulatie en metacognitie (Artelt et al., 2003). Onderzoeksartikelen van de laatste tien jaar hebben zich onder andere geconcentreerd op het leren van een diversiteit aan leerstrategieën. Hierbij richtten deze artikelen zich op het leren van domein specifieke of meer algemene leerstrategieën. Deze categorie omvat studies over het onderwijzen van probleemoplossing, metacognitieve strategieën, principes van wetenschappelijk onderzoek, hardop denken, ‘concept mapping’-organiseer-/structureermethoden, taalverwervingsstrategieën, fonologisch inzicht, leesstrategieën, schrijfstrategieën en tenslotte algemene leerstrategieën (Seidel en Steen, *ibid*).

Cognitieve activering door uitdagend en uitnodigen onderwijs. ‘Uitdaging’ wordt opgevat als instructie die zich richt op het activeren van het denkvermogen van leerlingen. Onderwijsaspecten zoals het denken op een hoger niveau, activering van het denkvermogen, diepgaand begrip van de leerstof, betekenisvolle contexten, authentieke instructie, relevantie van de leerstof, toepasselijke taal en taal op een hoog niveau, variatie in presentatiemethodes worden onderscheiden als factoren die een cognitieve uitdaging bieden voor leerlingen. Een curriculaire facet van cognitieve activering is dat aspiraties en leerdoelen worden gesteld op hogere niveaus van cognitieve verwerking, zoals transfer, probleemoplossen en evaluatie (vgl. de RTTI-methode; Drost en Verra, 2019).

Als het gaat om emotionele en motivationele ondersteuning kan worden begonnen met zich rekenschap te geven van het belang van leerkrachtverwachtingen. Het bekende verschijnsel van “Pygmalion in the classroom” illustreert de invloed van (te) vroege oordeelsvorming en stereotypering. Een op feitelijkheid gebaseerde houding tegenover dit verschijnsel pleit voor een kwalitatief verantwoord assessment, zodat a priori oordelen van docenten regelmatig getoetst kunnen worden.

In verband met het thema “soft skills”, signaleren Scheerens, Van der Werf en De Boer, (2020), dat sociaal-emotionele vaardigheden niet alleen als doel op zich kunnen worden beschouwd, maar ook als ondersteunend ten opzichte van het verwerven van cognitieve vaardigheden. Een intrigerend resultaat dat zij vermelden is namelijk dat in evaluaties van programma’s ter stimulering van sociaal-emotionele vaardigheden soms resultaten werden gevonden op cognitieve prestatie toetsen die vergelijkbaar waren met de effecten van programma’s die juist op verbetering van de cognitieve prestaties waren gericht. Een mogelijke verklaring van deze opmerkelijke resultaten kan worden gevonden in Bandura’s theorie over “self-efficacy” (Bandura, 1986), waarin aanmoediging door leerkrachten en vanuit de groep van medeleerlingen een rol speelt. Yeager en Walton, (2011) bespreken onderzoek waaruit zou blijken dat schijnbaar “kleine” en subtiele sociaal psychologische interventies een grote invloed kunnen hebben op gevoelens en motivatie van leerlingen rondom het presteren en taakgericht bezig zijn. Uiteindelijk wijzen ze op de wisselwerking van deze emotionele ondersteuning en de cognitieve leerervaringen van de leerlingen: “As we suggest, a key to understanding the long-lasting effects of social-psychological interventions is to understand how they interact with recursive processes already present in schools, such as the quality of students’ developing relationships with peers and teachers, their beliefs about their ability, and their acquisition of academic knowledge. It is by affecting self-reinforcing recursive processes that psychological interventions can cause lasting improvements in motivation and achievement even when the original treatment message has faded in salience” (*ibid*, p. 268). Een vruchtbare benadering om in het onderwijs aandacht te besteden aan sociaal-emotionele vaardigheden is daarom om dit

zoveel mogelijk “ingebed” in het normale onderwijs te laten plaatsvinden. In de volgende paragraaf bespreken we evaluatie en feedback als een effectiviteit bevorderende aanpak, die tevens een vruchtbare context is voor sociaal-emotioneel leren.

3.4.4 Evaluatie en feedback

Formatief toetsen heeft als doel het leerproces te verbeteren. Daartoe wordt regelmatig de voortgang van de leerlingen in het beheersen van leerstof vastgesteld en nagegaan hoe leerlingen zich ontwikkelen. Op grond van de bevindingen bepaalt de leraar hoe het leerproces het beste vorm kan krijgen. Cruciaal bij formatief toetsen is het geven van feedback. De term feedback komt voort uit de cybernetica en wordt ook gebruikt in de besturingstheorie (De Leeuw, 1990, p.126). De werking van een thermostaat wordt vaak gebruikt als voorbeeld om te demonstreren hoe feedback werkt. Wanneer op basis van een meting blijkt dat de kamertemperatuur beneden een bepaald niveau zakt, zorgt het regulerende mechanisme ervoor dat de verwarming aanslaat. De analogie met formatieve evaluatie in het onderwijs is correctief ingrijpen van de docent als de resultaten te laag zijn. De andere kant van de medaille is belonen van leerlingen bij goede resultaten en eventuele stimulering om op een hoger niveau verder te gaan.

Een veel gebruikte omschrijving van formatief evalueren (*formative assessment*) is die van Black & William (1998: 7). Zij omschrijven formatieve evaluatie als “alle activiteiten van docenten en leerlingen, die informatie geven die gebruikt kan worden als feedback om het onderwijs en de leeractiviteiten bij te stellen”. Nadere analyse laat zien dat er een onderscheid gemaakt kan worden in drie kernactiviteiten van formatief evalueren:

1. Feed-up en gegevensverzameling: bij feed-up gaat het om de vragen wat de leerdoelen zijn en welke succescriteria of mate van beheersing daarbij horen. De feed-up-fase is cruciaal bij formatieve evaluatie omdat het ontbreken van heldere doelen en succescriteria onduidelijkheid creëert. Daardoor kunnen de volgende twee kernactiviteiten, feedback en feed-forward, niet goed worden vormgegeven. Voor deze twee kernactiviteiten is informatie nodig over de leervorderingen van leerlingen. Deze informatie kan op meerdere manieren verzameld worden, niet alleen door het gebruik van toetsen, maar bijvoorbeeld ook met behulp van portfolio's, observaties, praktische opdrachten en leergesprekken.
2. Feedback: bij feedback staat de vraag centraal waar een leerling staat ten opzichte van de leerdoelen. Het is van belang dat leerlingen niet alleen een oordeel krijgen over wat zij goed en fout hebben gedaan, maar ook inzicht krijgen in waarom iets fout is en in mogelijke oplossingen en strategieën. Daarmee komen we bij de derde activiteit.
3. Feedforward en interventies: bij feed-forward gaat het om de vraag wat de volgende stap is die nodig is om het verschil tussen de gewenste situatie (feed-up) en de huidige situatie (feedback) te overbruggen. Interventies zijn concrete acties die daarbij uitgevoerd kunnen worden (Vgl. Oomens, Exalto en anderen (201).

In een bekende overzichtsstudie van Black & William, (1998) concluderen de auteurs dat formatieve evaluatie positief samenhangt met leerprestaties. In meta-analyses worden onderling sterk wisselende effectgroottes gevonden; vergelijk Hendriks, e.a. (2017). Een verklaring die wel gegeven is voor de betrekkelijk kleine effecten is dat het moeilijk is om het effect van formatieve evaluatie en feedback te onderscheiden van andere facetten van gestructureerd onderwijs. Dit blijkt bijvoorbeeld wanneer formatieve evaluatie onderzocht wordt in een context van “mastery learning”. Kluger & DeNisi (1996) hebben zich nader verdiept in de vraag wat feedback effectief maakt en concluderen daarbij dat zogenoemde instrumentele feedback beter werkt dan beoordelende feedback. Bij instrumentele

feedback wordt ingegaan op de oorzaken van het (niet bereiken) van een bepaald niveau, die tevens aangrijpingspunten bieden om tot verbetering te komen. Om instrumentele feedback te kunnen geven is het van belang dat er duidelijke doelstellingen en gewenste beheersingsniveaus zijn aangegeven en dat er tevens ideeën zijn over mechanismen en methoden om lacunes in het leren bij te werken.

Afgezien van de cognitieve werking van evaluatie en feedback zijn er ook emotionele en motivatiebepalende aspecten aan verbonden. Bij het stellen van doelen en beheersingsniveaus is het wenselijk dat de standaarden niet onhaalbaar zijn, maar ook weer niet te gemakkelijk (De Vos, 1989). Situaties die te maken hebben met terugkoppeling van resultaten zijn bij uitstek geschikt om sociaal-emotionele facetten van het leren op school aan de orde te stellen. Ervaringen van ongerustheid, teleurstelling, maar ook de voldoening van het bereikt hebben van doelen zijn levensechte oefensituaties voor sociaal-emotioneel leren.

Samenvattend kunnen we stellen dat evaluatie en feedback nauw verbonden zijn met directe instructie en geleid onderwijs. In Fig. 3.1 worden evaluatie en feedback gecategoriseerd als vormen van retroactieve regulering. Moderne onderwijsmethoden zijn vaak geheel gecentreerd om toetsen (vgl. Scheerens & Exalto, 2017, Drost & Verra, 2019)

3.5 Evidentie voor effectiviteit van de voornaamste categorieën van beïnvloedbare variabelen bij onderwijzen

In Scheerens, (2016) worden gegevens uit zo'n 70 meta-analyses gebruikt om een beeld te krijgen van de grootte van het effect van de variabelen in tabel 3.2. In het kader van deze studie zal alleen een samenvatting van dit materiaal worden weergegeven, voor details verwezen zij naar de hoofdstukken 8, 10 en 12 van deze publicatie en naar Scheerens et.al., 2007 en Scheerens, 2013. Omdat meta-analyses vaak onderlinge verschillen laten zien, is in de onderstaande tabel een vergelijking gemaakt tussen resultaten van "eigen" meta-analyses en die van andere auteurs, te weten Hattie, (2009), en Seidel & Shavelson, (2007).

Tabel 3.3 Resultaten van meta-analyses over effectieve instructie. De coëfficiënten zijn effectgroottes (coëfficiënt *d*). Ontleend aan Scheerens, 2016, p 245

	Scheerens et al., 2007	Seidel and Shavelson, 2007	Hattie, 2009
Time and OTL	.16	.08	.34
Classroom management	.20		.52
Structured teaching	.26	.06	.60
Teaching learning strategies	.44	.44	.70
Classroom climate	.24	.08	.54
Feedback & monitoring	.14	.02	.66

De resultaten van de meta-analyses verschillen aanmerkelijk. Het vermoeden bestaat dat dit voor een deel bepaald wordt door een sterke vertegenwoordiging van kleinschalige experimenten bij Hattie, terwijl in de meta-analyses van Scheerens et al. (2007) en Seidel & Shavelson (2007) vaker gebruik is gemaakt van grootschalige correlatieve studies.

Als een tweede voorbeeld van resultaten van meta-analyses over instructie-effectiviteit hieronder een overzichtstabel uit Hattie, (2008, p 243). De resultaten van Hattie zijn gebaseerd op grote aantallen meta-analyses. In de tabel worden onderwijsstrategieën waarbij de leerkracht een sterk activerende en structurerende rol speelt, vergeleken met aanpakken waarbij de leerkracht eerder als coach of begeleider optreedt.

Tabel 3.4: Effect sizes for teacher as activator and teacher as facilitator; geciteerd uit Hattie, 2009, p 243

Teacher as activator	d	Teacher as facilitator	d
Reciprocal teaching	.74	Simulation and gaming	.32
Feedback	.72	Inquiry-based teaching	.31
Teaching students self-verbalization	.67	Smaller class-sizes	.21
Meta-cognition strategies	.67	Individualized instruction	.20
Direct instruction	.59	Problem-based learning	.15
Mastery learning	.57	Different teaching for boys and girls	.12
Goals-challenging	.56	Web-based learning	.09
Frequent/effects of testing	.46	Whole language -reading	.06
Behavioral organizers	.41	Inductive teaching	.06
Average activator	.60	Average facilitator	.17

Bij de interpretatie van de betekenis van de effect-grootte coëfficiënten, sluiten we aan bij de standaard die is aangegeven door het Amerikaanse What Works Clearinghouse. Daarin worden effecten die groter zijn dan .25 als “onderwijskundig van betekenis” gezien. In de tekst nemen we een iets ruimere interpretatie en worden effectgroottes van .20 als onderwijskundig significant opgevat.

In de beschrijving van de hoofddimensies van effectieve instructie in de voorafgaande paragrafen blijkt het onderscheid tussen gestructureerd (docent geleid) versus open (of minimaal geleid) onderwijs belangrijk. Het gedachtengoed van het “constructivisme” heeft de populariteit van meer open, minder door de docent geleid onderwijs sterk doen toenemen. Echter het onderscheid bestaat al veel langer en wordt ook wel aangeduid met de termen traditioneel tegenover progressief onderwijs (vgl. eerdere verwijzing naar Peal, 2014). Ook in het overzicht dat ontleend is aan Hattie staat het onderscheid tussen gestructureerd onderwijs en “minder geleid” onderwijs centraal, met als uitkomst dat activerend, “geleid” onderwijs tot sterkere effecten leidt dan minder geleid onderwijs. Opmerkelijk is de plaats van meta-cognitieve strategieën aan de linkerkant van de tabel. Aandacht voor “leren leren” is eerder geassocieerd met het constructivisme en minder geleid onderwijs. Echter, zoals gezien bij de bespreking van “cognitieve activering”, bleek dat ook leerstrategieën op een gestructureerde manier onderwezen kunnen worden. De constatering dat de effectiviteit van “directe instructie” en gestructureerd onderwijs sterker onderbouwd is dan voor meer open onderwijsbenaderingen wordt verder benadrukt in de meta-analyse van Stockard et al. (2015). Van groot belang bij de interpretatie van deze onderzoeksuitkomst is dat er een overtuigende empirische onderbouwing is, op basis van de cognitieve theorie over de werking van het geheugen en informatieverwerking (Kirschner et al., 2006).

De empirische onderbouwing van “gelegenheid om te leren” en “onderwijstijd” is substantieel, maar laat door de bank genomen tamelijk bescheiden effecten zien. Voor de effectiviteit van het concept

“pedagogical content knowledge” lijkt vooralsnog nog geen sterke empirische basis te bestaan, ondanks positieve resultaten in enkele geavanceerde studies (Baumert et al., 2010). De overige modaliteiten van gestructureerd onderwijs, in de sfeer van klassenmanagement, klimaat en cognitieve en emotionele ondersteuning, worden in verschillende mate empirisch ondersteund. Er is een brede onderbouwing in het onderwijseffectiviteit onderzoek van het belang van een ordelijk klimaat; dit wordt ook ondersteund door uitkomsten in het kader van internationaal vergelijkend assessment onderzoek.

Klassenmanagement dat gericht is op effectieve leertijd en het bieden van “gelegenheid om te leren” (OTL) heeft een empirische ondersteuning die min of meer samenvalt met het effect van leertijd en OTL. Opmerkelijk is de empirische ondersteuning voor de effectiviteit van het aanleren van leerstrategieën (zie beide overzichten die gegeven zijn in de tabellen 3.3. en 3.4). Klassenmanagement dat breder gedefinieerd is en betrekking heeft op cognitieve en emotionele ondersteuning is onderzocht in een recente meta-analyse van Korpershoek et al. (2016). De gemiddelde effectgroottes die in deze meta-analyse gevonden werden bedroegen .17 voor cognitieve leerresultaten en .21 voor sociaal-emotionele uitkomsten. Specifieke werkvormen die in de meta-analyse van Scheerens et al. (2007) een klein, maar onderwijskundig relevant effect lieten zien (van rond de .20) waren samenwerkend leren en “peer tutoring”. In een recentere meta-analyses van Dietrichson et al. 2017 werden voor (adult) tutoring en coöperatief leren effect groottes van resp. .36 en .22 gevonden.

3.6 Hoe het beleid deze kennis over effectief onderwijzen zou kunnen gebruiken

Uit dit overzicht van concepten, empirische resultaten en theoretische principes ontstaat een beeld van effectief onderwijs. Dat beeld bevestigt het belang van onderdelen die klinken als open deuren, zoals het wijzen op het belang van goed opgeleide leerkrachten en een aanbod dat voldoende tijd en “gelegenheid” voor leerlingen biedt om te leren. Maar deze basisvoorwaarden worden nader uitgewerkt en gekwalificeerd.

Goede leerkrachten hebben een repertoire dat effectief onderwijs mogelijk maakt en in de tekst is daarbij verwezen naar “pedagogical content knowledge”; op te vatten als een synthese van kennis van inhouden en didactische strategieën, die passend zijn bij de behoeften van specifieke (groepen) leerlingen.

Een adequaat aanbod van onderwijs is in verband gebracht met concepten als “opportunity to learn” en een goed “uitgelijnd” curriculum, waarbij duidelijke onderwijsdoelstellingen en meting van onderwijsopbrengsten centrale componenten zijn.

De kern van effectieve instructie is getypeerd als directe instructie, een sterk empirisch gesteund concept, dat een bredere toepassing heeft dan feitelijke basiskennis, maar ook toepasbaar is in onderwijs gericht op hogere cognitieve operaties.

Andere ondersteunende condities die effectieve instructie bevorderen zijn besproken onder de noemer van klassenmanagement, een ordelijk klimaat en emotionele ondersteuning.

Evaluatie, assessment en feedback is afzonderlijk te noemen als een effectiviteit bevorderend mechanisme, als ankerpunt voor het leren van leerlingen en “bijsturen” van lesgeven.

Om nader te verduidelijken welke van de hierboven opgesomde componenten overeenkomen, dan wel contrasteren met het onderwijsbeleid van de laatste decennia, kan het volgende worden opgemerkt: In het concept van “curriculum alignment” wordt het assessment van eindtermen, in de vorm van examens en “high stakes” toetsen, meer centraal gesteld dan in recente benaderingen van curriculumvernieuwing, zoals “Onderwijs 2032” en “Curriculum.nu”.

Het gebruik van evaluatie, assessment en feedback komt overeen met het model van “opbrengstgericht werken”, dat enige tijd werd ondersteund en gefaciliteerd, maar dat inmiddels weer naar de achtergrond verdwenen is (zie ook hoofdstuk 5). Meer algemeen geldt dat evaluatie en assessment omstreden zijn in het Nederlandse onderwijs, zeker als het gaat om externe en objectiverende vormen van toetsing; ook al konden de nodige goede praktijken van formatief toets gebruik worden gesignaleerd (Vlg. Scheerens en Exalto, 2017 en Scheerens en anderen, 2019).

Niet zozeer als een manifestatie van overheidsbeleid, maar eerder als preferente cultuur in het georganiseerde onderwijsveld is de vraag te stellen naar het draagvlak voor het model van directe instructie in het Nederlandse onderwijs. Ondanks de kritiek van de Commissie Dijsselbloem uit 2008 zouden de huidige “eigenaarschapsideologie” en het benadrukken van extreme vormen van “gepersonaliseerd leren”, het onderwijs opnieuw in de richting van “open” onderwijs en ongeleide introductie van zelfregulering kunnen drijven. Tenslotte is de prioritering van vakgericht, op cognitieve doelstellingen gericht onderwijs in het geding. De betiteling hiervan door de Onderwijs Raad als een “smalle kijk op onderwijskwaliteit” relativeert het belang van deze inhoudelijke kern van het onderwijs ten onrechte. Een laatste significante bijdrage die in dezelfde richting gaat is de typering van onderwijskwaliteit als kwalificatie, socialisering en persoonsvorming, met de boodschap dat kwalificatie wel wat minder centraal zou mogen staan, terwijl de laatste twee meer accent zouden moeten krijgen (zie Scheerens et al, 2019, Scheerens, Van der Werf en De Boer, 2020). Kortom, de cultuur en “Zeitgeist” in het Nederlandse onderwijsklimaat anno 2021 lijkt niet onmiddellijk open te staan voor een wending naar vakgerichtheid, directe instructie en externe evaluatie.

Tegen deze achtergrond moeten de hefboomen worden gezien die de overheid heeft om effectief onderwijs, in de hierboven gegeven typering, te stimuleren. Deze worden hieronder samengevat.

Hefboomen op nationaal niveau om de dimensies van effectief onderwijzen te beïnvloeden zijn: maatregelen om het lerarentekort op te lossen, de initiële lerarenopleiding, de planning van het curriculum, her- en bijscholing, onderwijsinspectie en nationale toetsen en examens. In aanvulling hierop zou nog gedacht kunnen worden aan vormen van kwaliteitscontrole op het door de markt verzorgde aanbod aan onderwijsmethoden, zoals leerboeken en onderwijstechnologie. In Tabel 3.5 wordt een samenvattend overzicht gegeven.

Tabel 3.5: Hefbomen om de effectiviteit op 5 hoofddimensies te bevorderen

Effectiviteitsdimensies	Hefbomen
Voldoende gekwalificeerde leerkrachten	Aantrekkelijke arbeidsvoorwaarden voor leerkrachten, actief recruiteringsbeleid. Bevordering veelzijdig en goed geïntegreerd onderwijs repertoire bij aanstaande leerkrachten (“pedagogical content knowledge”)
Een kwalitatief hoogstaande lerarenopleiding	Richtlijnen die ervoor zorgen dat vakgericht onderwijs en een gestructureerde onderwijsaanpak een belangrijke plaats innemen
Een goed uitgelijnd curriculum	Expliciete nationale doelstellingen, eindtermen en standaarden, goed afgestemd op assessment en examens. Eventuele uitgewerkte leerlijnen en kwaliteitstandaarden voor leermiddelen, methoden en onderwijstechnologie
Gestructureerd onderwijs (didactiek)	Alleen indirect te beïnvloeden door initiële opleiding, de inhoud van verplichte toetsen en examens, de onderwijsinspectie, door de overheid gecertificeerde leermiddelen en methoden en eveneens gecertificeerde her- en bijscholing
Gestructureerd onderwijs (management en klimaat)	Idem
Goede voorzieningen voor evaluatie en feedback	Verplichte eindtoets en centraal examen

Bij dit schema is op te merken dat externe ondersteuning van scholen en leerkrachten bewust is weggegelaten. De voornaamste reden is dat wordt aangenomen dat in het Nederlandse systeem de overheid onvoldoende mogelijkheden heeft om de inhoud en kwaliteit van dergelijke ondersteuning te bepalen of te monitoren. Een andere reden is een overweging van doelmatigheid. In principe zouden de meer reguliere hefbomen in de figuur voldoende moeten zijn. En tenslotte zou het zelfs zo kunnen zijn dat deze niet te controleren ondersteuning contraproductief werkt, door het omhelzen van modieuze en niet op evidentie gebaseerde onderwijsvisies en benaderingen.

In dit stadium ga ik niet verder in op de haalbaarheid van deze toepassing van hefbomen, gezien vigerende normatieve kaders, ideologische posities en het besturend vermogen van de overheid. We komen daarop terug, na effectiviteit bevorderende factoren op meso- en macroniveau te hebben behandeld in de volgende twee hoofdstukken.

4 WAT WERKT ER OP HET MESO-NIVEAU (SCHOOLORGANISATIE EN SCHOOLLEIDERSCHAP)

Op basis van het Multi-level model dat besproken is in Hoofdstuk 2 (Fig. 2.1) moet het schoolniveau gezien worden als intermediair tussen het centrale niveau en het primaire proces op het microniveau. Enerzijds ontvangt de school als organisatie impulsen vanuit de centrale overheid en anderzijds wordt de functie van de school als organisatie vooral gezien als het faciliteren en stimuleren van effectieve instructie op het microniveau. In dit hoofdstuk wordt in de eerste plaats een overzicht gegeven van de voornaamste variabelen die, vanuit dit perspectief, gebruikt zijn in het onderzoek naar schooleffectiviteit. Vervolgens wordt ingegaan op theoretische modellen van de school en de plaats die de variabelen uit het effectiviteitsonderzoek daarin hebben. Daarbij wordt in het bijzonder stilgestaan bij concepten van “onderwijskundig leiderschap” en “de lerende organisatie”. Zoals op het microniveau gestructureerd versus meer “open” onderwijs een belangrijk thema is, is er op het schoolniveau een vergelijkbare koppel van onderscheiden prioriteiten, namelijk tussen rationeel empirische modellen en modellen waarin de menselijke verhoudingen (human relations) centraal worden gesteld. Het hoofdstuk wordt afgesloten met een impressie van de resultaten van empirisch school effectiviteitsonderzoek en implicaties voor overheidsbeleid.

4.1 Overzicht van variabelen geadresseerd in school effectiviteitsonderzoek

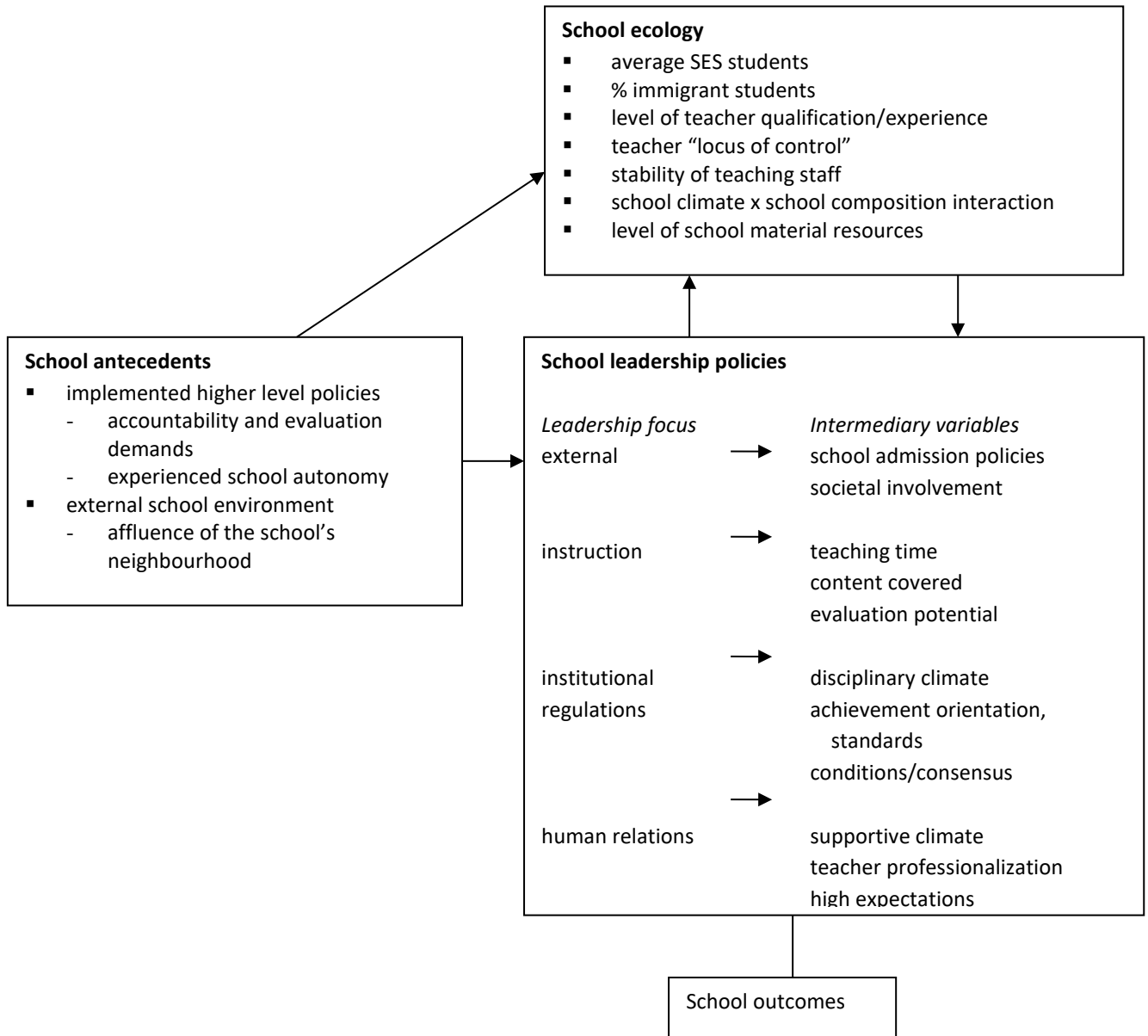
Over de lijst van schoolkenmerken die als effectiviteit bevorderend worden gezien bestaat al decennialang aanzienlijke consensus (vgl. onder meer overzichtsstudies van Purkey and Smith, 1983, en Scheerens, 1992). In een recentere review studie uit 2014 (Reynolds et al., 2014), werd deze consistentie opnieuw bevestigd (Scheerens, 2014), vergelijk het overzicht in Tabel 4.1.

Tabel 4.1: effectiveness enhancing conditions referred to in the review study by Reynolds et al. (2014, as summarized by Scheerens (2014))

Educational Effectiveness Research
Effective Leadership
Academic focus
A positive orderly climate
High expectations
Monitoring progress
Parental involvement
Effective teaching (time)
Staff professional development
Pupil involvement

Bij dit overzicht is op te merken dat er sprake is van enerzijds pure organisatiefactoren, zoals leiderschap, ouderbetrokkenheid en professionele ontwikkeling van de staf, en anderzijds factoren die zowel betekenis hebben op het microniveau van onderwijzen en lesgeven als op het niveau van de schoolorganisatie. Dit geldt voor de factoren die staan voor een academische gerichtheid, een positief en ordelijk klimaat, het bijhouden van de vorderingen van leerlingen (monitoring), effectieve onderwijstijd, “hoge verwachtingen” en de betrokkenheid van de leerlingen. In het onderzoek worden

deze variabelen soms geaggregeerd tot het schoolniveau, dan weer opgenomen in multi-level analyses en bij andere gelegenheden opgevat als “het beleid van de school” om bijvoorbeeld monitoring op klasniveau te bevorderen. Enkele eveneens vaak onderzocht variabelen die niet in het overzicht in Fig. 4.1 zijn opgenomen zijn: samenwerking en consensus tussen leerkrachten, differentiatie in de klas, en “gelegenheid om te Leren”. Een meer gedetailleerd overzicht wordt gegeven in Figuur 4.1 en Tabel 4.2.



Figuur 4.1: Variabelen in een model van het functioneren van de school; uit Scheerens, 2016, p. 100.

Tabel 4.2: School variabelen, geciteerd uit Scheerens, 2016, p. 100-101.

School antecedents	School ecology	School leadership, policies, and organization
<p><u>Implemented higher level policies</u></p> <ul style="list-style-type: none"> ▪ accountability and evaluation demands ▪ experienced school autonomy ▪ national curriculum frames <p><u>External school environment</u></p> <ul style="list-style-type: none"> ▪ affluence of the school neighborhood 	<ul style="list-style-type: none"> - student composition (e.g. school average SES) - percentage of students from immigrant background; - percentage of students with a special education profile - teacher composition (e.g. the average qualification, experience and locus of control scores of the teachers) - stability of teaching staff over the last 5 years - school climate x school composition interaction - level of school material resourcing - level of extra-curricular activities of the school 	<ul style="list-style-type: none"> - Leadership focus; specifically the degree of instruction oriented leadership; - Achievement orientation/ high expectations - Teaching time - Quality of school curriculum, opportunity to learn; - Coordination, cooperation and consensus among staff; - A safe and orderly climate, supporting positively experienced interactions - Parental involvement - Opportunities for professional development of teachers - School admission policies. - Evaluation, feedback and monitoring - Degree to which teachers participate in decision making

De overzichten in Figuur 4.1 en Tabel 4.2 volgen de structuur en de opbouw van het model uit Fig. 2.1. In de kolom “voorwaardelijke condities”, die betrekking hebben op de macrocontext en de directe omgeving van de school, staan onder meer maatregelen en structuren die door de overheid bepaald worden. Binnen de dimensie “school ecologie” spelen compositievariabelen, die gezien worden als gegeven maar tevens beperkt beïnvloedbaar door de school, een belangrijke rol. Een variabelencategorie die niet vaak wordt onderzocht zijn compositiekenmerken van de groep leerkrachten. Van deze variabele wordt aangenomen dat zij medebepalend is voor het optimisme en de prestatiegerichtheid op school, als onderdeel van de informele cultuur van de organisatie. De direct door de school beïnvloedbare variabelen, waarvan wordt aangenomen dat zij effectieve instructie bevorderen, zijn neergezet als facetten van leiderschap en het creëren van gunstige organisatorische randvoorwaarden.

4.2 Theorieën over de school als organisatie

In deze paragraaf wordt een overzicht “in vogelvlucht” gegeven van de belangrijkste theorieën over scholen als organisaties, die directe relevantie hebben voor een beter begrip van schooleffectiviteit. In een volgende paragraaf wordt dan iets meer in detail ingegaan op twee van de zes modellen, namelijk onderwijskundig leiderschap, als onderdeel van het model van de doel rationele school en “de school als lerende organisatie”. Een overzicht wordt gegeven in de onderstaande tabel 4.3.

Tabel 4.3: School models and key variables

School model	Key variables
<i>Fend's new theory of the school</i>	Legal context and formal arrangements (e.g. curricula, mandatory tests) Dynamic interactions (e.g. implemented curriculum, matching of school rules and classroom culture)
<i>Professional bureaucracy</i>	Initial teacher training, professional development
<i>Effective schools model</i>	Achievement, orientation, high expectations Educational leadership Consensus and cooperation of staff Curriculum quality and opportunity to learn School climate Evaluative potential Parental involvement Effective learning time
<i>School as learning organizations</i>	Professional development Evaluation and feedback
<i>School culture as a product of self-organization</i>	Characteristics of teachers Student composition effects Teacher composition effects Complex interactions among teachers
<i>Schools as high reliability organizations</i>	Externally developed curricula and lesson plans Frequent monitoring Institutional climate of schools

Fend's nieuwe theorie van de school

Fend, (1980, 2006) beschrijft de school als een onderdeel van een multiniveau onderwijsstelsel dat enerzijds gereguleerd wordt door institutionele normen en formele voorschriften en anderzijds functioneert op basis van dynamische interacties die een zekere ruimte bieden voor zelfontwikkeling en "zelf-referentie". De twee theoretische referentiekaders achter deze twee dimensies zijn enerzijds “institutionalisme” en anderzijds de systeemdymanica (waarbij hij zich vooral laat inspireren door het werk van de Duitse socioloog Luhmann). De beide dimensies laten opnieuw bipolaire extremen zien, enerzijds expliciete regels, die deels extern zijn vastgesteld, anderzijds vrije interactie. De neiging tot decentralisatie en deregulering van de laatste decennia hebben in veel landen tot autonomievergroting geleid. Tegelijkertijd vestigt de theorie de aandacht op de legitimiteit en ook de daadwerkelijke mogelijkheid voor de overheid om door regelgeving invloed uit te oefenen op scholen. Een ander aspect

van deze tweedeling is de relatieve vrijheid en speelruimte die er kan zijn om extern aanbod op een eigen manier te implementeren.

De school als professionele bureaucratie

Dit organisatieconcept is ontwikkeld door Mintzberg (1979). De voornaamste kenmerken van de school als professionele bureaucratie zijn als volgt:

- De interne cohesie van de organisatie hang voornamelijk af van de standaardisatie van de vaardigheden van de functionarissen, leerkrachten in dit geval, die gebaseerd is op een lange en gespecialiseerde opleiding.
- Een grote mate van professionele autonomie van de leerkrachten, waarbij de loyaliteit aan de eigen organisatie wedijvert met de loyaliteit aan de professie en de "client".
- Leiderschap en management in de professionele bureaucratie zijn vooral administratief en grijpen niet in op het primaire proces van het lesgeven. Schoolleiders spelen hierbij een ondergeschikte, hooguit ondersteunende rol.
- Een betrekkelijk onderontwikkelde belangstelling voor de externe omgeving; een aanname hierbij is dat, hoewel complex, de omgeving in wezen stabiel is.
- De technologie in de professionele bureaucratie is enerzijds te kenmerken als een "goed gevulde gereedschapskist", maar wordt tegelijkertijd uitgedaagd door de behoefte om aan te sluiten bij de steeds wisselende omstandigheden in het werken met cliënten (in dit geval leerlingen).
- Er is weinig bereidheid tot verandering bij de professionals en er is zelfs verzet tegen rationalisering van het werk (bijvoorbeeld in de vorm van extern vastgestelde leerlijnen) en tegen het volgen en beoordelen van de prestaties.
- Rekrutering van personeel wordt gezien als de belangrijkste stuurmaatregel om de organisatie te veranderen en te verbeteren. Binnen de professie als zodanig is bijstelling van de initiële opleiding van de professionals de voornaamste hefboom tot verandering.

De professionele autonomie van leerkrachten en het weinig ophebben met externe regulering geven de professionele bureaucratie gemakkelijk een fragmentarisch karakter. Door Weick (1976) worden scholen en universiteiten als prototypes van "losjes gekoppelde systemen" gezien. Hoewel sommige van de nog te bespreken "modernere" visies op het functioneren van scholen als organisaties meer de richting op gaan van management, monitoring en rationalisering, zijn bepaalde basiskenmerken van blijvende betekenis. Het principe van de professionele autonomie van leerkrachten is nog steeds actueel, evenals de moeizame zoektocht naar acceptatie van coördinatie, leiderschap en externe verantwoording. Volgens de theorie van de professionele bureaucratie zou men beducht moeten zijn voor over-regulatie, die bovendien ten koste gaat van de ingebouwde efficiëntie van autonome leerkrachten.

Prescriptieve interpretatie van het schooleffectiviteitsmodel: de rationeel geleide school

Onderwijskundig leiderschap

Schooleffectiviteitsonderzoek is een vorm van empirisch onderzoek. De resultaten ervan, bijvoorbeeld zoals die zijn samengevat in Fig. 4.1, zijn kenmerken die positief geassocieerd zijn met leerresultaten. Deze informatie biedt aangrijpingspunten voor op verbetering gerichte activiteiten, waarbij de schoolleider een belangrijke plaats wordt toegekend. Anders gezegd, de uitkomsten van het onderwijseffectiviteitsonderzoek, zijn te zien als hefboomen voor verbetering. Het type leiderschap dat hierop gericht is wordt aangeduid met termen als onderwijskundig of instructiegericht leiderschap. Onderwijskundig leiderschap is uiteindelijk gericht op betere leerresultaten van leerlingen. Dit gebeurt voor een deel indirect, doordat de schoolleider condities creëert die leerkrachten ondersteunen bij het geven van effectieve instructie (vgl. hfst 3). Anderzijds is onderwijskundig leiderschap gericht op bredere

kenmerken van de school, die het functioneren van de school als geheel ten goede komen. Voor het gemak duiden we hier de eerste vorm aan als “instructiegericht” leiderschap en de tweede als “organisatiegericht leiderschap”.

De voornaamste kenmerken van “instructie-gericht” leiderschap zijn:

- Een tijdsbesteding van de schoolleider waarin ruimte is voor onderwijskundige, naast administratieve, verplichtingen.
- De schoolleider als “meta-bestuurder” van het onderwijs op het niveau van de klas of leergroep.
- De schoolleiders als toezichthouder op de kwaliteit van het onderwijskundig handelen van leerkrachten.
- De schoolleider als ondersteuner van taakgerichte teams.
- De schoolleider als initiator en ondersteuner van op verdere professionalisering van de leerkrachten gerichte activiteiten.

Het begrip meta-besturing vraagt enige toelichting. Hiermee wordt niet bedoeld dat de schoolleider de leerkrachten voortdurend over de schouders kijkt. Met respect voor de professionele autonomie van leerkrachten betekent het eerder een zichtbaar engagement met het werk in de klas, een betrokkenheid bij belangrijke beslissingen over de keuze van doestellingen, methoden en vormen van assessment.

Kenmerken van een breder organisatiegericht leiderschap zijn:

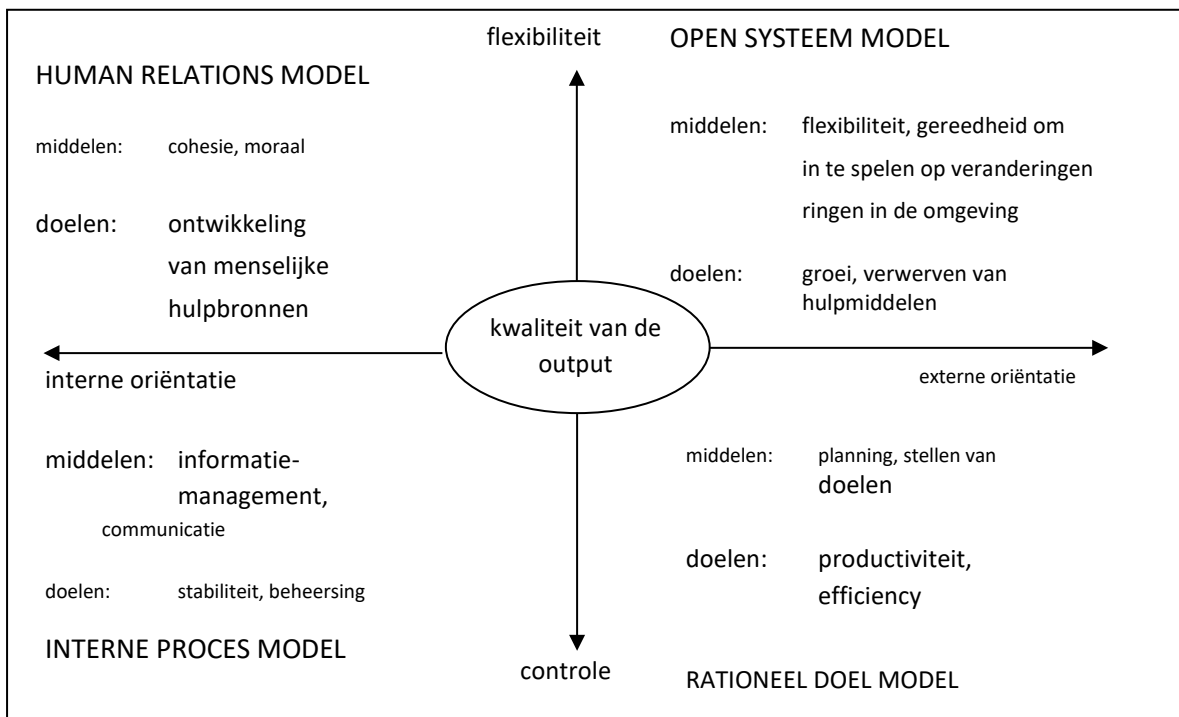
- het vervullen van een helder uitgewerkte en zichtbare leiderschapsfunctie;
- het zorgdragen voor informatievoorziening;
- het onderhouden van contacten met extern betrokken organisaties en partijen (zogenoemde “stakeholders”);
- het zorgdragen voor inspraak en participerende besluitvorming;
- het stimuleren van een prestatiegericht beleid;
- het stimuleren van samenwerking tussen leerkrachten en gezamenlijk commitment aan een prestatiegerichte missie van de school (vgl. Leithwood and Montgomery, 1982, p. 334).

Onderwijskundig leiderschap ingebed in een breder model van organisatie-effectiviteit

De differentiatie in accenten in schoolleiderschap krijgt meer reliëf door aan te sluiten bij een breder model van organisatie-effectiviteit, ontwikkeld door Quinn en Rohrbaugh (1983).

In dit model worden meerdere effectiviteitsperspectieven onderscheiden, die bepaald worden door een accentuering van verschillende soorten doelen en oriëntaties. Quinn en Rohrbaugh onderscheiden vier effectiviteitsperspectieven, die bepaald worden door twee onafhankelijke dimensies. De eerste dimensie heeft betrekking op de interne dan wel externe oriëntatie van de organisatie. De tweede dimensie heeft te maken met de mate van structurering: oriëntatie op controle en beheersing enerzijds versus gerichtheid op flexibiliteit en verandering anderzijds.

Door deze twee dimensies met elkaar te kruisen ontstaan vier kwadranten. Elk ervan representeert volgens Quinn en Rohrbaugh een effectiviteitsmodel.



Figuur 4.2 : Typologie van effectiviteitsmodellen, naar: Quinn & Rohrbaugh, 1983

Bij verdere analyse van deze vier modellen valt op dat zij onderling van nogal uiteenlopende aard zijn. In het "open systeem model" wordt in feite de vraag centraal gesteld of de organisatie is toegerust voor het op een juiste wijze kiezen van doelen. De oriëntatie is hier meer op gericht op "doing the right things" dan op "doing things right". Het model is het meest relevant in situaties waarin de organisatie in een dynamische omgeving verkeert. De oriëntatie op doelkeuze binnen dit model is in verband te brengen met het beeld van de lerende organisatie en met het concept van "double-loop" learning, waarbij een organisatie de eigen doelen ter discussie stelt (vgl. Morgan, 1986, p. 88). "Doing things right" is veeleer de oriëntatie van het rationele doelmodel, waarin een spaarzame en trefzekere keuze van middelen ter bereiking van de primaire productiedoelen van de organisatie centraal staat. Het "formatief" evalueren van alternatieve middelen is een voorbeeld van "single loop learning", zoals dat in de conceptualisering van de lerende organisatie wordt behandeld. De beide modellen aan de linkerkant van figuur 3, het "human relations model" en het "interne procesmodel" hebben ontwikkeling van "human resources", respectievelijk stabiliteit en beheersing als doelen. Men zou deze zaken echter ook kunnen opvatten als ondersteunende condities of middelen ten dienste van de productiviteit van de organisatie. Op deze wijze wordt er een middel-doel ordening tussen de vier opvattingen van organisatorische effectiviteit aangebracht.

Schooleffectiviteitsonderzoek is vooral onder te brengen in het rationele doelmodel. De kernvraag is welke door de school beheersbare condities een positieve invloed hebben op de leerresultaten, nadat gecontroleerd is voor, zo mogelijk, de prestaties aan het begin van een bepaalde leerperiode en andere achtergrondkenmerken van leerlingen (Scheerens, 1992; Scheerens en Bosker, 1997). Kwaliteitsindicatoren die gebruikt kunnen worden om de verschillende kwadranten van het model van Quinn en Rohrbaugh te meten, zijn samengevat in tabel 4.4; belangrijke procesindicatoren afkomstig uit schooleffectiviteitsonderzoek zijn opgenomen onder de noemer van het rationele doelmodel.

Tabel 4.4: *Proces indicatoren van schoolkwaliteit op basis van het model van Quinn en Rohrbaugh en schooleffectiviteitsonderzoek. Naar: Scheerens, Luyten en Van Ravens, (2011, p 8)*

<p>Human relations model</p> <ul style="list-style-type: none"> - onderling respect - participerende besluitvorming - professionele interactie - feedback op werkgedrag en prestaties - gelegenheid om vaardigheden toe te passen - overeenkomst tussen persoonlijke en organisatie doelen - professionele ontwikkeling 	<p>Open Systeem model</p> <ul style="list-style-type: none"> - ondernemerschap - collegialiteit - openstaan voor zelfevaluatie en leren - school marketing - ouderparticipatie - belang van “boundary spanning positions” - samenwerking met externe ondersteuners van vernieuwing - kerngetallen over instroom en uitstroom van leerlingen - adequate materiele middelen en voorzieningen
<p>Interne proces model</p> <ul style="list-style-type: none"> - gebruik van planningsdocumenten - toepassing van Management Informatie Systemen - formalisering van taken en aanstellingen - continuïteit en stabiliteit in personeel en leiding - goed onderling afgestemde leergangen - het bijhouden van aanwezigheid en absentie - registratie van lesuitval 	<p>Rationeel doel model (schooleffectiviteitsonderzoek)</p> <ul style="list-style-type: none"> - onderwijskundig leiderschap - prestatiegerichtheid - het bijhouden van de vorderingen van de leerlingen - leertijd - correspondentie leeraanbod en toetsing (opportunity to learn) <p>(bredere verzameling van leerdoelen)</p> <ul style="list-style-type: none"> - lesgeven in teams - individualisering en differentiatie - tijd voor sociale, emotionele, creatieve en morele ontwikkeling - “leren leren”

Op basis van deze verdere differentiatie van accenten in schoolleiderschap wordt nog duidelijker dat er een tweedeling kan worden gemaakt tussen leiderschap dat gericht is op het primaire proces (vgl het rationele doelmodel) en leiderschap dat gericht is op ondersteunende, secundaire organisatieprocessen. De verhandeling over de school als professionele bureaucratie maakt duidelijk dat “instructiegericht leiderschap” als het meest ingrijpend in de traditionele verhoudingen moet worden gezien. De transactielasten voor schoolleiders bij het realiseren van instructiegericht leiderschap moeten als zwaarder worden ingeschat dan ondersteunend leiderschap op secundaire organisatiefuncties.

Indirecte effectmodellen bij schoolleiderschapsonderzoek

Het model van Quinn en Rohrbaugh maakt duidelijk dat er verschillende accenten en prioriteiten in schoolleiderschap worden gelegd. Bekend is bijvoorbeeld ook het onderscheid tussen “transactioneel” en “transformatief” leiderschap. Transactioneel leiderschap komt overeen met onderwijskundig leiderschap, geassocieerd met het rationele doelmodel en outputgerichtheid. Transformatief leiderschap is gericht op verandering in bredere zin, met een sterk accent op “human relations”-modellen en samenwerking met externe ondersteuners van vernieuwing. Zoals eerder in deze paragraaf

is aangegeven kan de effectiviteitsvraag voor wat betreft schoolleiderschap enerzijds worden gezien als een directe relatie tussen schoolleidersgedrag en leerresultaten, maar is vooral een indirecte relatie, gemedieerd door andere condities van effectieve instructie. In de onderstaande figuur wordt de focus in leiderschap, geïnspireerd op het Quinn en Rorbauch-model in verband gebracht met condities van effectieve instructie. In zogenoemde indirecte effectmodellen zijn deze condities geplaatst als intermediaire variabelen. Dit is weergegeven in de onderstaande tabel, die ook onderdeel is van het overkoepelende model, weergegeven in Fig. 4.1.

Tabel.4.5 Intermediaire variabelen en accent in schoolleiderschap (Naar Kruger en Scheerens, 2012, p.18)

Leadership focus	Intermediary variables
external	school admission policies societal involvement
instruction	teaching time content covered evaluation potential
institutional	disciplinary climate
regulations	achievement orientation, standards conditions/consensus
human relations	supportive climate teacher professionalization high expectations participative decision making

In paragraaf 4.3 wordt ingegaan op de empirische ondersteuning voor dit soort indirecte modellen van schoolleiderseffectiviteit.

Vanuit het perspectief van onderwijseffectiviteit is onderwijskundig leiderschap, als manifestatie van het rationele doelmodel, de meest logische keuze. Tegelijkertijd kunnen oriëntaties die te maken hebben met externe relaties, formele regelgeving en “human relations” als ondersteunende condities worden meegenomen. Gezien de spanning tussen de professionele autonomie van de leerkracht (vgl de verhandeling over de professionele bureaucratie) en onderwijskundig leiderschap dat deze autonomie niet ongemoeid laat, bestaat het risico dat het schoolleiderschap zich verliest in condities die verder verwijderd zijn van het primaire proces. Dit kan worden gezien als een vorm van “goal displacement” en organisatieontwikkeling om de organisatieontwikkeling.

De school als lerende organisatie

Het concept van de lerende organisatie gaat over het leren van functionarissen binnen organisaties (in ons geval leraren) en over het leren van organisaties als geheel. Een lerende organisatie heeft beide typen van leren geïnstitutionaliseerd. Wat dit precies betekent zal in de opbouw van deze paragraaf geleidelijk duidelijk worden gemaakt. Ik begin met het verduidelijken van organisatie-leren en bespreek dan later of de institutionalisering van dit leren zou moeten leiden tot een bepaald organisatiebeeld, dat bijvoorbeeld vergelijkbaar zou zijn met het concept van de school als professionele bureaucratie.

Organisatie-leren wordt op drie manieren gedefinieerd:

- a) Als de som van het leren van alle functionarissen in de organisatie.
- b) In de zin van het bevorderen en consolideren van de effectiviteit van de organisatie (te typeren als enkelslag leren; “single loop learning”, zoals gedefinieerd door Argyris & Schoen (1974)

- c) In de zin van het versterken van de responsiviteit van de organisatie, hetgeen niet alleen betrekking heeft op middelen en methoden, maar ook kan leiden tot een revisie van de doelen van de organisatie (dubbelslag leren).

Ad a) *Het leren van individuele functionarissen*. Initiële training wordt in het organisatiemodel van de school als lerende organisatie als fundamenteel gezien. Daarnaast is er de laatste decennia toenemende aandacht voor her- en bijscholing van leerkrachten en ook voor “training on the job” en continue professionele ontwikkeling. Een basale invulling van “de school als lerende organisatie” is dan ook dat de individuele inspanningen tot verdere professionele ontwikkeling gecoördineerd verlopen. “Training on the job” bestaat niet alleen in de vorm van externe ondersteuning, maar ook volledig intern, onder noemers als “professionele consultatie”, “peer learning” en “professionele leergemeenschappen”. In sommige interpretaties van “de school als lerende organisatie” is deze samenwerking een van de kernelementen, waarbij “leren” gelijk wordt gesteld aan collegiale uitwisseling van ervaringen (Vgl. Kools & Stoll, 2016).

Ad b) *Organisatie-leren als enkelslag leren*. De concepten enkel- en dubbelslag leren vormen de kern van de theoretische basis van de lerende organisatie. Enkelslagleren berust op de capaciteit om fouten te detecteren en te corrigeren als het gaat om het nastreven van een gegeven doelstelling (Morgan, 1986, p 88). Op zijn beurt moet enkelslag leren gezien worden in het kader van het cybernetisch principe, dat staat voor “stuurmanschap”. Op basis van dit principe worden processen van zelfregulatie van organismen en organisaties gezien in termen van informatie-uitwisseling die betrekking hebben op negatieve feedback (zie ook de beschrijving van cybernetica in Hfst 3). Leren wordt gekarakteriseerd als een geleidelijke bijstelling en aanpassing van gedrag, waarbij voortdurend gecorrigeerd wordt op fouten en suboptimale oplossingen. In een sociale context berusten “goed” en “fout” op overeenkomsten en normen. Wanneer deze nader zijn gepreciseerd, moeten ze worden opgevat als de na te streven opbrengsten van de organisatie. Bij enkelslag-leren worden deze doelstellingen als gegeven beschouwd en concentreert men zich op de optimale selectie van middelen en technologie om deze doelstellingen te bereiken. Dit instrumentele perspectief sluit perfect aan bij het paradigma van school- en onderwijs-effectiviteitsonderzoek, waarin gebruik wordt gemaakt van de wetenschappelijke methode om uit te vinden welke organisatie- en instructiecondities het meest doeltreffend zijn om de onderwijsdoelstellingen te bereiken. In theorie kan ook het van dag tot dag runnen van een organisatie worden gezien als zijnde geleid door deze instrumentele benadering, al gebeurt dit op een minder gestileerde en geformaliseerde manier dan bij wetenschappelijk onderzoek. In de praktijk blijken er heel wat hindernissen en barrières te zijn voor zo’n benadering. Deze zullen we verderop bespreken. Samengevat: enkelslag leren benadrukt de behoefte aan informatie die kan leiden tot een geleidelijke verbetering van primaire en secundaire organisatieprocessen om de kerndoelen van het primaire proces te realiseren.

Ad c) *Dubbelslag leren*. Bij dubbelslag leren gaat het niet alleen om het vinden van de beste methoden bij gegeven doelen, maar om het analyseren van die doelen en operationele standaarden zelf. Het basismotief is gebaseerd op een “open systeem” visie op organisaties. In het overkoepelende systeem model, dat besproken is in Hfst. 2, is deze visie opgenomen door de incorporatie van een omgevingscomponent, evenals het in aanmerking nemen van relaties tussen eenheden op verschillend niveau en de omgeving. De capaciteit van een organisatie om zich rekenschap te geven van eventuele veranderde eisen en verwachtingen vanuit de omgeving werd aangeduid met de term “responsiviteit”. Met name twee soorten facetten van de omgeving van scholen worden hierbij in aanmerking genomen, enerzijds veranderingen in de voorspelbaarheid van de omgeving (bijvoorbeeld demografische veranderingen in de omgeving waar de school zijn leerlingen uit

betreft) en anderzijds veranderingen in de technologie, beschikbaar voor het primaire proces (bijvoorbeeld toepassingen van artificiële intelligentie in de onderwijstechnologie). Het gegeven dat organisaties zich op deze manier rekenschap geven van omgevingscondities sluit aan bij een belangrijke organisatietheorie, de zogenoemde contingentie theorie.

Een belangrijk leerstuk uit de theorie van organisatie-leren zijn de zogenoemde “barrières” bij zowel enkelslag als -dubbelslag leren (Argyris & Schoen, 1978). Beide vormen van organisatie-leren zijn rationele modellen, die in theorie leiden tot optimale oplossingen. Barrières treden op wanneer organisatieleden, organisatieonderdelen of individuele leden van de organisatie in feite andere doelstellingen nastreven. Zo kan het zijn dat organisatieleden het beleid van de leiding niet onderschrijven en wanneer dit leidt tot wederzijds onbegrip, kan dit weer leiden tot “inhibitory loops”, waardoor de toestand verergerd wordt. Leerproblemen kunnen ook het gevolg zijn van onvoldoende kennis en “technische” tekortkomingen van de functionarissen. Remedies om hieraan tegemoet te komen worden gezocht in overleg, samenwerking en “professionele ontwikkeling”. In onderwijsorganisaties zijn weerstanden tegen de assessment – en evaluatie technologie een gevoelig thema, dat enkelslag-leren al in de kiem kan smoren.

Om de betekenis van de theorie over organisatie-leren en de lerende organisatie voor de effectiviteit van scholen nader op zijn merites te bezien, beperken we ons tot enkele “fundamentals”.

Om te beginnen stellen wij vast dat het “cybernetisch principe” (evaluatie en feedback) de kern vormt van enkelslag-leren en daarmee ook van dubbelslag-leren.

In de tweede plaats hangt het belang van dubbelslag-leren voor scholen af van een inschatting over de dynamiek van de omgeving en ook in welke omgevingsfactoren die dynamiek vooral zit. Vaak wordt erg gemakkelijk gezegd dat, omdat de wereld zo snel verandert, de school ook snel zou moeten veranderen.⁵ Een hypothese zou kunnen zijn dat meer diversiteit qua leerling-instroom belangrijker is dan veranderingen in de georganiseerde omgeving van de school. Een verder gezichtspunt is dat omgevingsdynamiek het resultaat kan zijn van een overproductie van beleid en steeds wisselende prioriteiten van de overheid. Vanuit ons systemisch perspectief ligt het dan voor de hand om deze problematiek op systeemniveau te leggen en niet bij de school.

In de derde plaats is het begrijpelijk dat pogingen om de samenwerking tussen functionarissen onderling en functionarissen en de schoolleiding te verbeteren worden ingezet om enkelslag-leren te ondersteunen en weerstanden tegen de toepassing van het cybernetisch principe te verminderen. In gangbare interpretaties van de school als lerende organisatie lijkt samenwerking soms als doel op zich te worden gezien (zie bijvoorbeeld Kools en Stoll, 2016).

In de vierde plaats kan men de betekenis van het op schoolniveau institutionaliseren van dubbelslag-leren ter discussie stellen. Hiermee staat of valt de toegevoegde waarde van het spreken van een “lerende organisatie”. De kern van het verhaal is institutionalisering van een goede invulling van enkelslag leren, hetgeen neerkomt op een adequate toepassing van het cybernetisch principe, dat weer bepaald wordt door een goede praktijk van assessment, evaluatie en feedback.

In de vijfde plaats tenslotte is het risico van de metafoor van de lerende organisatie overproductie van sociaal agogische activiteit, waarbij interne en externe samenwerkingsactiviteiten doel op zich zijn geworden en de kern van evaluatie en feedback naar de achtergrond is verdwenen. Dit opblazen van het concept blijkt bijvoorbeeld uit de volgende opsomming van Kools en Stoll (2016, p 10) van kenmerken van de lerende organisatie.

⁵ Meanwhile, schools are urged to learn fast and teachers are urged to become knowledge workers in order to deal effectively with the growing pressures of a rapidly changing environment (Schleicher, 2018)

- developing and sharing a vision centred on the learning of all students
- creating and supporting continuous learning opportunities for all staff
- promoting team learning and collaboration among all staff
- establishing a culture of inquiry, innovation and exploration
- embedding systems for collecting and exchanging knowledge and learning
- learning with and from the external environment and larger learning system
- modelling and growing learning leadership

Een schoolcultuur die de effectiviteit bevordert.

In het schooleffectiviteitsonderzoek is het schoolklimaat vaak onderzocht als een van de hefboomen van prestatieverbetering. Daarmee wordt dan gedoeld op facetten van de algemene sfeer op school die te maken hebben met orde en regelmaat, veiligheid en vertrouwen, goede onderlinge betrekkingen en prestatiegerichtheid. Met de schoolcultuur wordt iets overeenkomstigs bedoeld. Vaak wordt geconstateerd dat de termen ongeveer hetzelfde betekenen (vergelijk b.v. Connolly & Kruse, 2019). We kiezen hier voor de term cultuur, omdat de fundamenteën van dit begrip beter zijn ontwikkeld in de organisatie-literatuur. De schoolcultuur kan worden gedefinieerd als de “elementaire assumpties, normen en waarden, en culturele artefacten die gedeeld worden door de schoolstaf en die hun functioneren op school beïnvloeden” (Maslowski, 2001, p. 8 en 9). Schein (1985) onderscheidt drie lagen in de organisatiecultuur. De onderste laag bestaat uit elementaire aannamen, die naar zijn mening, de essentie vormen. Het gaat daarbij om overtuigingen die door de leden van de organisatie gedeeld worden, maar zodanig vanzelfsprekend worden gevonden dat men er zich in het dagelijks functioneren nauwelijks meer van bewust is. Het gaat om vijf groepen van elementaire aannamen namelijk over: de relatie van de organisatie met de omgeving, de aard van de werkelijkheid en wat als waar wordt beschouwd, de menselijke natuur, de aard van de menselijke bezigheden en de aard van intermenselijke relaties. Zo wordt bijvoorbeeld bij aannamen over de menselijke natuur gedacht aan de vraag of de mens van nature goed of slecht is, of mensen grotendeels bepaald worden door hun aangeboren eigenschappen en of mensen plooibaar en verbeterbaar zijn. De middelste laag in Schein’s hiërarchie beziet de organisatiecultuur als een geheel van waarden en normen. Waarden, zoals een positieve waardering van samenwerken en respect worden vaak vertaald in normen en gedragsregels, die op dit niveau nog meer het karakter hebben van “ongeschreven regels”. Dergelijke normen kunnen bijvoorbeeld betrekking hebben op de kleding die leraren aantrekken en wat leerkrachten al dan niet geacht worden te doen op het vlak van professionele ontwikkeling. De hoogste en meest zichtbare laag in de typering van de organisatiecultuur door Schein bestaat uit meer zichtbare zaken zoals symbolen, vaste gedragspatronen, verhalen en mythes. Een voorbeeld zou een charismatische leerkracht kunnen zijn, die als rolmodel gaat fungeren. Hoewel dit niveau in de buurt komt van expliciet beleid en formele afspraken worden ze nog steeds gezien als zich ontwikkelend uit de dagelijkse gang van zaken en langzaam ingeslepen gewoonten. Deze hoogste en meest zichtbare laag in het cultuurmodel van Schein komt overeen met wat door Argyris en Shoemaker (1978, p. 11) wordt omschreven als “theory in use”, misschien vrij te vertalen als de dagelijkse praktijktheorie. Hierbij gaat het om gewoonten en gebruiken, die uit het gedrag van de leden van de organisatie opgemaakt kunnen worden. Daarnaast onderscheiden deze auteurs “the espoused theory”, als zijnde de officiële interpretatie van hoe de dingen op een bepaalde school gaan en behoren te gaan.

In al deze concepten van de schoolcultuur ligt het accent op regelmatigheden die zich voordoen en gedeeltelijk spontaan naar boven komen, met andere woorden eerder “emergent” dan “malleable” zijn; dat wil zeggen spontaan gegroeid in plaats van expliciet gepland gedrag. In het school

effectiviteitsonderzoek is vaak gekeken naar prestatiegerichtheid in de zin van specifiek geplande maatregelen en ordelijkheid en veiligheid op basis van schoolregels. Het concept "hoge verwachtingen" van de prestaties van leerlingen is weer wat lastiger als planbaar voor te stellen en zou kunnen samenhangen met persoonlijkheidskenmerken van individuele leerkrachten.

De vraag kan worden gesteld wat nu de meerwaarde is van het in aanmerking nemen van de "diepere" lagen van de organisatiecultuur. Naar mijn mening is dat in de eerste plaats dat een dergelijke analyse inzicht geeft in de begrensde maakbaarheid van dit verschijnsel. In de tweede plaats bestaan er intrigerende theoretische verklaringen van "emergente" regelmatigheden in het functioneren van organisaties. De complexiteitstheorie beschrijft processen van zelforganisatie die voortkomen uit ingewikkelde interacties (vgl. bv Stacey et al., 2000). Pogingen om dit soort processen zodanig te begrijpen dat er een bijdrage wordt geleverd aan schoolverbetering zijn tot nu toe vrij beperkt, evenals de concrete resultaten ervan. (Vlg. Scheerens, 2008, O' Day, 2002, Daily, Molenaar & Carrier, 2011). Persoonlijk vind ik de veronderstelling dat compositie-effecten die gebaseerd zijn op persoonlijkheidseigenschappen van leraren van invloed zijn op de schoolcultuur interessant en empirisch onderzoekbaar. Naast deze speculaties over systeemdynamica en "emergente" condities zijn er ook organisatie-theorieën die zich juist op formele regels en voorschriften richten en daarin de belangrijkste basis zien voor effectiviteitsverbetering. Een voorbeeld daarvan wordt besproken in de volgende paragraaf.

Het schoolklimaat is eerder met de meer gedragsmatige facetten van de schoolcultuur te associëren. Dit geldt in het bijzonder voor gedragsregels en het handhaven van discipline. In het vorige hoofdstuk werd al gewezen op het belang van een ordelijk klimaat in de klas. Ondersteuning hiervan op het niveau van de school speelt hierbij een belangrijke rol. Het ligt voor de hand dat regels en omgangsvormen schoolbreed worden vastgelegd.

Scholen als "High reliability organizations"

De ontwikkeling van de latente schoolcultuur als een proces van interne endogene ontwikkelingen kapitaliseert op de *praxis* dimensie uit de theorie van de school van Fend, die in een vorige paragraaf besproken is. De institutionele theorie (vgl Meyer & Rowan, 1977) benadrukt juist de andere dimensie uit Fend's model, namelijk die van institutionele normen en formele regulering. Deze nadruk op formele regulering komt overeen met de klassieke theorie van de bureaucratie. McMeeking (2003) gebruikt deze theorie als inspiratiebron voor zijn stelling dat het "institutioneel klimaat" van de school gezien moet worden als de sleutel voor schooleffectiviteit. Hij stelt dat "het institutionele klimaat van de school dat gebaseerd is op formele regels, informele regels, mechanismen om beide soort regels te handhaven, duidelijke doelstellingen en een atmosfeer die bepaald wordt door vertrouwen en samenwerking, een sterke invloed heeft op de performance van een school" (Ibid, p1). Een organisatie-model dat hiermee verwantschap vertoont is dat van "high reliability organizations", dat onder andere geïnspireerd is op de manier waarop de besturing van kernreactoren en luchtverkeersleidingssystemen functioneert. Stringfield (1995) brengt dit organisatie-model in verband met succesvolle schoolverbeteringsprojecten in de VS. Een van die projecten was het "Calvert-Barclay schoolproject", waarin een sterk gestructureerd en traditioneel academisch georiënteerd programma dat werd toegepast op een privé school (private school) met succes werd toegepast in een school in een achterstandswijk van een grote stad. "Hoog betrouwbare organisaties" hebben de volgende kenmerken:

- De notie dat mislukking als rampzalig wordt beschouwd.
- Absolute duidelijkheid van de doelstellingen bij de staf, evenals een sterk commitment aan de primaire doelstelling van de organisatie (nl optimalisering van leerprestaties).
- Het gebruik van "standard operating procedures", zoals uitgeschreven lesprotocollen.

- Hoge prioriteit voor de rekrutering van leerkrachten en zorg voor intensieve her- en bijscholing waar nodig.
- Voorzieningen gericht op het blootleggen van fouten, (monitoringsystemen).
- Veel aandacht voor de evaluatie van de prestaties, niet alleen gericht op beoordeling, maar ook op verbetering van processen.
- Monitoring wordt als wederzijds gezien, zonder contraproductief verlies aan vertrouwen en autonomie.
- Alertheid voor verstoringen en terugval in functioneren (de notie dat aanvankelijk kleine fouten tot groter falen kunnen leiden).
- Een hiërarchische structuur, die toch ruimte laat aan collegiale besluitvorming, ook in crisissituaties.

Ondanks de kritiek op het nut van rationele planning en “mechanistische” structuren voor onderwijsorganisaties (bijvoorbeeld Lotto & Clark, 1986) laten deze voorbeelden zien dat er plaats is voor geformaliseerde en sterk gestructureerde onderwijsprogramma’s in een gedisciplineerde cultuur en met “alle neuzen dezelfde kant op”.

4.3 Empirische resultaten van schooleffectiviteitsonderzoek

In de onderstaande tabel worden de uitkomsten van 4 overzichtsstudies weergegeven. De tabel geeft een overzicht van de effectgroottes van een aantal kernvariabelen, opgenomen in Fig. 4.1. De variabelen zijn geordend op basis van de effectgroottes die werden gevonden in de meta-analyse van Marzano, 2003, in de eerste kolom.

Tabel 4.6: Rank ordering of school effectiveness variables according to the average effect sizes (*d*-coefficient) reported in three reviews/meta-analyses.

	Marzano, 2003	Scheerens et al. 2007	Hattie, 2009	Creemers & Kyriakides, 2008
Opportunity to learn	.88	.30	.39 *	.30
Instruction time	.39	.30	.38	
Monitoring	.30	.12	.64	.36
Achievement pressure	.27	.28	.43 **	
Parental involvement	.26	.18	.50	
School climate	.22	.26	.34	.24
Homework		.14		
School leadership	.10	.10	.36	.14
Cooperation	.06	.04	.18 ***	.32
Differentiation ⁶		.03		

*) operationalized as “enrichment programs for gifted children”;

**) operationalized as “teacher expectations”;

***) operationalized as “team teaching”.

⁶ In de tabel is differentiatie op schoolniveau gedefinieerd. In Scheerens et al. (2007) is adaptief onderwijs ook op klasniveau onderzocht, en bleek daar een effect-grootte van .07 te hebben. Hattie (2008) rapporteert een effectgrootte van geïndividualiseerd onderwijs van .20.

Een deel van deze variabelen verwijst naar processen die op het microniveau gedefinieerd zijn (gelegenheid om te leren, instructie tijd, huiswerk, prestatiedruk en differentiatie). Hier zijn deze variabelen opgevat als schoolvariabelen, in de zin dat ze staan voor het schoolbeleid en de schoolbrede toepassing van deze factoren. Een ander deel van de variabelen zijn gedefinieerd op het schoolniveau (schoolleiderschap, ouderbetrokkenheid en samenwerking tussen leraren). De effectgroottes zijn betrekkelijk klein, met uitzondering van de coëfficiënt die door Marzano wordt gerapporteerd voor gelegenheid om te leren. Nadere analyse wijst uit dat deze uitkomst is gebaseerd op een viertal oudere onderzoeken, met daarbij een “uitbijter” (een extreem hoge waarde). De effectgroottes voor schoolleiderschap, samenwerking tussen leerkrachten en differentiatie zijn zelfs bijna verwaarloosbaar. Hetzelfde beeld, van kleine effecten van de schoolvariabelen gerelateerd aan relevante onderwijsfactoren en van verwaarloosbare effecten van “zuivere” schoolorganisatie variabelen blijkt uit de resultaten van Hattie (2009, p106-107).

Het is een opmerkelijke uitkomst dat schoolorganisatiefactoren die vooral de aandacht krijgen in de schoolverbeteringsliteratuur (Fullan, Stoll), namelijk leiderschap en samenwerking tussen docenten, zulke kleine effecten laten zien in het schooleffectiviteitsonderzoek. We exploreren dit thema verder door nader in te gaan op uitkomsten van onderzoek naar de effecten van leiderschap, samenwerking tussen leerkrachten en organisatie-leren.

Aanvullende gegevens over de effecten van schoolleiderschap

In het voorafgaande hebben we resultaten gepresenteerd, die variantie in schoolleidersgedrag hebben gerelateerd aan leerresultaten van leerlingen, met bijvoorbeeld als hypothese dat schoolleiders, die meer onderwijskundig leiderschap manifesteren, betere resultaten laten zien dan schoolleiders die dit in mindere mate doen. De onderzoeksvraag is hier: welke kenmerken heeft effectief leiderschap. Een daar in zekere zin aan voorafgaande vraag is; in hoeverre het een verschil maakt of een school geleid wordt door schoolleider A of schoolleider B. Hier is de onderzoeksvraag naar de mate waarin de variantie in leerresultaten wordt bepaald door de factor schoolleider. Eenzelfde type vraag kwam ter sprake in Hoofdstuk 3, waarin het ging om de bijdrage van de variantie tussen individuele leerkrachten aan de totale variantie in leerresultaten. Er zijn enkele studies van economen, die de vraag “hoeveel doet de schoolleider er toe” hebben proberen te beantwoorden (Branch, Hanushek & Rivkin, 2013, en Coelli & Green, 2012). Het analytische probleem dat moet worden opgelost is het scheiden van het schoolleiderschap van het “overall” schooleffect. Dit is alleen mogelijk door te kijken naar data waarin transities van schoolleiders zijn opgenomen, zodat kan worden nagegaan of een schoolleider consistent geassocieerd kan worden met goede leerprestaties. Coelli & Green concluderen dat de impact van de schoolleider sterk afhangt van de tijd van aanwezigheid op een bepaalde school.

Een relatief recente ontwikkeling in het onderzoek naar schoolleiderschapseffecten is het analyseren van zogenoemde indirecte effectmodellen. Hierbij zijn drie typen van variabelen in het geding. In de eerste plaats variabelen die een bepaald type leiderschap weergeven, in de tweede plaats effectvariabelen en in de derde plaats intermediaire processen. Met behulp van pad-analyse technieken, zoals LISREL is het mogelijk om met deze variabelen zowel directe, indirecte en totale effecten te schatten (waarbij de totale effecten de som zijn van het directe en de indirecte effecten). Een voorbeeld van het basismodel is: onderwijskundig leiderschap > effectieve leertijd > leerresultaten. Inhoudelijk interessant is de aard van de mediërende variabelen. Hendriks & Scheerens, (2013) onderzochten intermediaire effectmodellen van 15 studies, met in totaal 35 replicaties. De gemiddelde totale effectgrootte was .06. Slechts in een enkel geval waren er indirecte effecten van enige substantie (de studie van Heck & Moryani, 2010). In de 15 onderzoeken werd een breed scala aan intermediaire variabelen onderzocht variërend van de academische focus van het onderwijs tot de aansturing van het

instructieprogramma en het schoolklimaat. Hoewel het onderzoek van Heck & Moryani, (2010) een duidelijk effect van het instructieprogramma als intermediaire variabele kon vaststellen, is dit een te smalle basis om te kunnen worden gezien als een overtuigende basis voor de hypothese dat onderwijskundig leiderschap effectief is door aansturing van het instructie-proces. Een recente meta-analyse van Liebowitz & Porter, (2020) rapporteert een gemiddelde effectgrootte van instructiegericht leiderschap van .11, en van.16 voor interactiegericht leiderschap.

De bevinding, gedocumenteerd in tabel 4.5, van betrekkelijke bescheiden effecten van schoolleiderschap, wordt bevestigd in enkele overzichten van meta-analyses in Hendriks en Scheerens, (2014, tabellen 2.1 en 2.5). Daar komt bij dat een deel van de meta-analyses en de onderliggende onderzoeken veel methodologische tekortkomingen bevat (ibid). Een treffend voorbeeld is het onderzoek van Day e.a. (2009). Het onderzoek werd gepresenteerd als een “mixed methods” studie, een combinatie van kwantitatieve en kwalitatieve methoden. De conclusie luidde dat er sterke verbanden zijn aangetoond tussen schoolleiderschap en prestaties van leerlingen. In werkelijkheid leverde het kwantitatieve deel van de studie een effectgrootte van .01 op, terwijl in het kwalitatieve deel van het onderzoek zeer positieve zelfbeoordelingen de overhand hadden. Het is een meer algemeen verschijnsel bij het onderzoek naar schoolverbetering dat er in case-studies en beschrijvingen van “goede praktijken” hoog wordt opgegeven over het belang van de schoolleider bij het initiëren en ondersteunen van op verbetering gerichte samenwerkingsactiviteiten. Bekend zijn de verhalen over schoolleiders die in betrekkelijk korte tijd een laag presterende school weten om te turnen naar een sterschool. Het probleem met deze gegevens is dat er meestal geen resultaatmetingen zijn gedaan en dat generalisatie problematisch is.

De effectiviteit van de “lerende organisatie” nader bezien

Uit de beschrijving van de theorie van de lerende organisatie in paragraaf 4.2 kwam naar voren dat het cybernetisch principe (evaluatie en terugkoppeling) de harde kern is van het concept van de lerende organisatie. Als ondersteunend mechanisme hierbij kan taakgerichte samenwerking tussen functionarissen onderling en overeenstemming met de leiding genoemd worden. Het overzicht in tabel 4.5 laat een klein tot middelmatig effect zien van evaluatie/monitoring en assessment en een bijna te verwaarlozen effect van samenwerking. In deze paragraaf kijken we nog wat nader naar de empirische evidentie wat het betreft de effectiviteit van beide processen (interne evaluatie en samenwerking op school).

In de onderstaande tabel, geciteerd uit het proefschrift van Maria Hendriks, wordt een overzicht gegeven van resultaten van meta-analyses naar de effecten van evaluatie en assessment op leerprestaties (Hendriks, 2014, p. 143)

Tabel 4.7 Overview of earlier meta-analyses on the effects of evaluation and assessment on student achievement; geciteerd uit Hendriks, 2014, 143).

Authors	Headings of evaluation, assessment or feedback concepts in studies included	Estimated mean effect size	Number of studies included
<i>Evaluation</i>			
Fuchs & Fuchs (1986)	Providing formative evaluation to teachers	$r = .33^*$	21
Scheerens, Luyten, Steen & Luyten-de Thouars (2007)	Monitoring at school and class level	$r = .06$	43
Kyriakides, Creemers, Antoniou & Demetriou (2010)	Student assessment (school level)	$r = .18$	12
	Evaluation of school policy on teaching and actions taken for improving teaching practice (school level)	$r = .13$	6
Kingston & Nash (2011)	Formative assessment	$r = .10$	13
Kyriakides, Christoforou & Charalambos (2013)	Assessment (class level)	$r = .34$	27
<i>Assessment</i>			
Bangert-Drowns, Kulik, & Kulik (1991)	Frequency of classroom testing	$r = .11$	35
Kim (2005)	Performance assessment	$r = .17$	148
Seidel & Shavelson (2007)	Evaluation of learning (class level)	$r = .02$	10
Hattie (2009)	Frequency or effects of testing	$r = .17$	569
<i>Feedback</i>			
Kluger & DeNisi (1996)	Feedback	$r = .19$	131
Hattie & Timperley (2007)	Feedback	$r = .35$	196

* De gemiddelde effectgrootte van de meta-analyses is bepaald op basis van de Fischer-Z transformatie van de correlatiecoëfficiënt. De r-waarden zijn gelijk aan de helft van de d-coëfficiënt.

De gemiddelde effectgrootte over deze 11 meta-analyses is gelijk aan .40 (d-coëfficiënt) De eigen door Hendriks uitgevoerde meta-analyse leverde aanmerkelijk lagere gemiddelde effecten op. Voor “evaluatie op schoolniveau” werd een effect grootte van .14 vastgesteld en voor “evaluatie op klasniveau” eveneens .14. Het meer algemene gebruik van “assessment” leverde een effectgrootte van .02 op. Alles bij elkaar laten de verzamelde meta-analyses een klein tot middelmatig effect van evaluatie/monitoring en assessment zien (+/- .30).

De effecten van ‘samenwerking’ tussen docenten, zoals samengevat in tabel 4.5 zijn verwaarloosbaar tot klein. In onderzoek naar zogenoemde ‘professionele leergemeenschappen’ wordt samenwerking nader gekwalificeerd, op een manier die beter recht doet aan het bredere concept van de lerende organisatie. In het volgende citaat ontleend aan Toole & Louis, (2002, p. 249), wordt dit concept uiteengelegd in vier componenten, namelijk *reflectieve dialoog, docenten die elkaar wederzijds feedback*

geven over hun lesgeven, een gezamenlijke focus en een collectieve gerichtheid op het leren van leerlingen (geciteerd uit Lomos, Hofman & Bosker, 2011a, p 124). Lomos et al. (ibid) voerden een meta-analyse uit die was gebaseerd op een vijftal onderzoeken en vonden een gemiddelde effectgrootte van .25, wat zij interpreteren als een klein maar relevant effect. In een onderzoek op Nederlandse scholen voor voortgezet onderwijs gingen dezelfde auteurs na of vaksecties functioneerden als professionele leergemeenschappen en wat het effect was op leerprestaties. Wat deze laatste vraag betreft werd een overall effect van .24 gevonden. Verder werd geconstateerd dat het positieve effect van professionele leergemeenschappen verklaard werd door de component “gerichtheid op leren”, terwijl er geen effecten werden gevonden voor componenten als “reflectieve dialoog” en “samenwerking” (Lomos, Hofman & Bosker, 2011b). Zoals ook uit het literatuuroverzicht in beide studies blijkt, heeft het begrip “professionele leergemeenschappen” geen algemeen aanvaarde begripsbepaling. In de traditie van schoolorganisatie-ontwikkeling en sociaal-agogische benaderingen van schoolverandering (vgl het werk van auteurs als Miles en Fullan) wordt samenwerking eerder als doel op zich gezien en niet als instrumenteel ten opzichte van leren en leerprestaties. Het Groningse onderzoek bevestigt dat samenwerking effectief is bij een taak- en op prestatiesgerichte invulling.

4.4 Discussie en aangrijpingspunten voor beleid

Wat werkt er op het niveau van de school? Ik vat eerst samen welke antwoorden voortkomen uit de beschreven organisatiemodellen, kijk daarna nog eens naar de empirische evidentie en kom daarna met een antwoord over effectieve en minder effectieve hefboomen. Tenslotte wordt weer stilgestaan bij de vraag wat de overheid kan doen.

Modellen

Het model van de professionele bureaucratie benadrukt de autonomie van leerkrachten. Aan dat principe wordt in recentere modellen van de school als organisatie niet echt getornd, al wordt erkend dat onderwijskundig leiderschap impact kan hebben door voorwaarden voor effectieve instructie te faciliteren. Het model van de school als lerende organisatie bekrachtigt het belang van evaluatie en feedback (cybernetisch principe) als een van de sturende mechanismen in de rationeel geleide school. Daarnaast kapitaliseert de lerende organisatie op samenwerken (professionele leergemeenschappen). De theoretische onderbouwing van de schoolcultuur wijst enerzijds op minder zichtbare en controleerbare interacties en compositie-effecten, maar anderzijds op het bestaan van formele regelgeving om een veilig en ordelijk schoolklimaat te creëren. Beide dimensies, vrije interacties en formele regelgeving, staan ook centraal in Fend’s model van de school. Het model van de school als “high reliability organization” gaat het verst in formele regulering en de voorstructurering van het primaire proces.

Evidentie

De uitkomsten van schooleffectiviteitsonderzoek zijn in drie categorieën te verdelen:

1. Factoren die in feite de aggregatie tot het schoolniveau zijn van condities van effectieve instructie (onderwijstijd, schoolcurriculum en gelegenheid om te leren, huiswerkbeleid, prestatiegerichtheid en differentiatie).
2. Factoren die staan voor de directe ondersteuning van het primaire proces, zoals onderwijskundig leiderschap, regulering met het oog op veiligheid en discipline, het evaluatie- en assessmentbeleid van de school en taakgerichte samenwerking.

3. Algemener schoolorganisatie kenmerken, zoals samenwerking tussen leerkrachten en leiderschap dat meer gericht is op externe relaties.

Als we kijken naar de empirische evidentie van elk van deze categorieën, dan moet in de eerste plaats worden opgemerkt dat alle effecten van de variabelen op schoolniveau betrekkelijk bescheiden zijn. De resultaten op elk van deze categorieën zijn ruwweg⁷ in de orde van grootte van .30 voor de eerste categorie (aggregatie effectieve instructiekenmerken), .20 voor de tweede categorie (waaronder onderwijskundig leiderschap) en minder dan .10 voor transformationeel leiderschap en samenwerking.

Nadere interpretatie

De “logica” van deze uitkomsten is dat door de school beheersbare factoren die dichter bij het primaire onderwijsproces liggen er meer toe doen dan betrekkelijk algemene secundaire organisatiecondities. “So far, so good”, zou men kunnen zeggen, alles werkt een beetje, dus laten we al deze hefboomen om het onderwijs op peil te houden en te verbeteren maar aangrijpen. Vanuit een doelmatigheidsperspectief is zo’n conclusie discutabel. Hoe belangrijk zijn investeringen in secundaire condities (categorie 3) nu eigenlijk en moet er geen oog zijn voor eventuele contraproductieve effecten van al die aandacht voor organisatieleren, samenwerking en ook de overdracht van steeds meer functies naar de leerkracht, vergelijk concepten als “teacher leadership” en de leerkracht als onderzoeker? We treffen deze concentratie op secundaire organisatiefactoren aan in de eerder geciteerde brede invulling die Kools en Stoll (2016) geven aan het concept van de “lerende organisatie”, maar ook in de McKinsey filosofie over gradaties in excellentie van onderwijsstelsels. Volgens die beschouwing is voor de hogere niveaus van excellentie “peer based learning” van leerkrachten en “new pedagogy” nodig (vgl Mourshed et al., 2010). Een derde voorbeeld is de uitdijning van het schoolleiderschapsconcept, van instructie-leiderschap naar transactie-leiderschap en verder naar organisatieleiderschap. Vergelijk het volgende citaat:

While earlier research emphasized the importance of a knowledge base of effective instruction and instructional leadership, it is clear that principals also need skills in organizational management to unleash the potential of... teachers... through the removal of barriers and creation and refinement of conditions that influence school culture” (Donley & States, 2020, p 17). De empirische onderbouwing van het tweede deel van deze bewering is nog weinig overtuigend. Wel sluit de bewering van Donley & States aan bij de “human relations” benadering van organisatieverandering van auteurs als Miles en Fullan).

Bij de beoordeling van de effectiviteit van schoolleiderschapsbenaderingen moeten verder twee perspectieven in het oog worden gehouden. Volgens het eerste perspectief zijn we op zoek naar factoren die het verschil maken en die duidelijk effect hebben op de verbetering van leerprestaties. Volgens het andere, meer beschrijvende perspectief, kunnen we vaststellen dat het takenpakket van schoolleiders breed en divers is en is aannemelijk dat externe contacten, “people management” en andere niet direct op het primaire procesgerichte activiteiten wel degelijk belangrijke randcondities zijn, zonder dat moet worden aangenomen dat ze echt het verschil maken bij schoolverbetering. De in dit hoofdstuk besproken organisatiemodellen suggereren dit zeker niet.

Wat het beleid kan doen

De hefboomen voor beleid die aan het eind van het vorige hoofdstuk zijn geformuleerd en die betrekking hadden op verhoging van de effectiviteit op het microniveau zijn evenzeer van belang voor het mesoniveau. De uitkomst dat schoolkenmerken, die als geaggregeerde instructiekenmerken kunnen

⁷ Afzonderlijke condities binnen de categorieën vertonen soms aanzienlijke verschillen, verder zijn er verschillen in beschikbaarheid van empirische gegevens.

worden gezien, op schoolniveau de meeste invloed hebben onderstreept deze conclusie. Het belang van een eindtoets in het basisonderwijs en centrale examens in het voortgezet onderwijs heeft op dit niveau een aanvullende betekenis, omdat output-toetsing scholen stimuleert tot een effectieve en efficiënte focus op het primaire onderwijsproces en een rem zet op “overproductie” op het vlak van secundaire organisatiefuncties. Het beleid, maar ook de sector als geheel, zou een kritische houding moeten innemen als het gaat om de groei van “management overhead” en ondersteuning op brede organisatieontwikkeling.

In meer positieve zin zou kunnen worden bijgedragen aan de totstandkoming van een functieprofiel van schoolleiders dat in lijn is met de evidentie over effectief onderwijs en onderwijskundig leiderschap. In het onderstaande overzicht zijn hefboomen die de centrale overheid zou kunnen gebruiken om aspecten van een effectieve en efficiënte schoolorganisatie te stimuleren samengevat.

Tabel 4.8 Hefboomen om een effectieve en efficiënte schoolorganisatie te stimuleren

Effectiviteitsdimensies	Hefboomen
Een schoolorganisatie die effectieve instructie ondersteunt	Indirect door de overheid te stimuleren door accountability toetsing en een daarop gerichte onderwijsinspectie
Een schoolorganisatie die onderhevig is aan output-gerichte accountability, en die dit gebruikt om een prestatiegerichte focus te hebben	In stand houden eindexamens, hernieuwde uniformering van de eindtoets basisonderwijs en het bijhouden van prestatie ranglijsten van scholen
Een schoolorganisatie die goed gebruik maakt van evaluatie en feedback-processen	Positief neven-effect van output gerichte accountability, in de zin dat externe accountability interne accountability stimuleert (Vgl. Carnoy et al. 2003)
Een schoolorganisatie die efficiënt is in leiderschapsfuncties (onderwijskundig leiderschap) en zuinig is met investeringen op secundaire organisatie factoren	In de Nederlandse context moeilijk aanstuurbaar door de Rijks- overheid.
Een schoolorganisatie die gebruikt maakt van formele regels om een ordelijk klimaat te stimuleren	Idem

Zoals uit het overzicht al blijkt zijn de mogelijkheden voor de overheid om direct te sturen op de effectiviteit verhogende condities op schoolniveau beperkt. Hier zou verandering in aangebracht kunnen worden door de *lump sum* financiering te vervangen door doel subsidies, maar hiervoor lijkt in de sector geen draagvlak te bestaan. Uit een review van de OECD blijkt verder dat externe evaluatie van schoolleiders en leerkrachten in Nederland een witte plek is (OECD, 2014). Dit is overigens niet anders in veel andere OECD, landen, al zijn er landen (Oostenrijk, sommige “Laender” in Duitsland) waar de onderwijsinspectie ook individuele leerkrachten beoordeelt. Tegelijkertijd is er een omvangrijke private en semi-publieke dienstverlening, die professionalisering en ondersteuning aanbiedt. In dit aanbod spelen multi—nationals en “private equity” een steeds grotere rol. Het is aan te bevelen om na te gaan of de overheid hier via regulering meer grip op zou kunnen krijgen.

5 WAT WERKT OP SYSTEEMNIVEAU?

5.1 Inleiding

In dit hoofdstuk wordt ingezoomd op het centrale systeemniveau, zoals weergegeven in Fig. 2.1. De indeling van het hoofdstuk volgt de structuur die ook in de vorige twee hoofdstukken is gebruikt. Om te beginnen wordt een overzicht gegeven van de variabelen die object van onderzoek zijn geweest op systeemniveau. In dit hoofdstuk gebeurt dat in 2 twee stappen. Eerst wordt een brede beschrijving van systeemindicatoren gegeven, die zowel voor meer beschrijvend als op effectiviteit gericht onderzoek zijn gebruikt. Daarna volgt een toespitsing op de factoren die gebruikt zijn in eerder effectiviteitsonderzoek en wordt nader ingegaan op de conceptualisering en theoretische onderbouwing van een beperkt aantal kernvariabelen die betrekking hebben op de structurele kenmerken van onderwijsstelsels. Van deze wordt de empirische evidentie besproken op basis van onderwijseffectiviteitsonderzoek en internationaal vergelijkend assessment onderzoek. De resultaten van dit laatste type onderzoek worden in meer detail besproken in de Annex bij dit rapport. Een aparte plaats in dit hoofdstuk hebben facetten van “lopend beleid”, zoals de jaarlijkse onderwijsbegroting en landelijke kwaliteitsverbeteringsprogramma’s. Gezien de diversiteit en omvang van dit gebied wordt alleen een meer impressionistische beschrijving en interpretatie gegeven. Het hoofdstuk wordt, net als de voorafgaande hoofdstukken, afgesloten met een afsluitende discussie en resumé van de belangrijkste hefboomen voor het beleid.

5.2 Overzicht van variabelen op systeemniveau

5.2.1 Een brede beschrijving van systeemindicatoren

De overzichten in de tabellen 5.1 tot en met 5.4 zijn ontleend aan (Scheerens, 2011), en zijn onder andere gebaseerd op activiteiten in het kader van het OECD-project over internationale onderwijsindicatoren INES en de UNESCO (Scheerens, 2007).

Tabel 5.1: Voorbeelden van manipuleerbare procesindicatoren op systeemniveau

Procesindicatoren gedefinieerd op het niveau van het nationaal onderwijssysteem

- Curriculum kenmerken, zoals het aantal lessen per vak, en het totaal aantal lessen per jaar, per vak, en per schoolcategorie (po en vo)
- De gelegenheid om te leren (“opportunity to learn”), in de vorm van een beoordeling door deskundigen van de overlap tussen eindtoetsen en nationale curricula
- De locus van besluitvorming in het onderwijs, per onderwijsniveau (deze indicator geeft aan op welk bestuursniveau de besluitvormingsbevoegdheid ligt op subdomeinen van het onderwijs – curriculum, personeelsmanagement, en de inzet van financiële en materiele “resources”)
- Autonomie van de school (deze indicator maakt deel uit van het concept van de locus van besluitvorming)
- Prestatienormen of standaarden per niveau (bijvoorbeeld targets voor het slagingspercentage of het percentage leerlingen dat een bepaald prestatieniveau bereikt of overtreft)
- Of er al dan niet formele toetsing plaatsvindt aan het einde van elke schoolcategorie

- De mate waarin er sprake is van “streaming” of differentiatie in schoolsoorten (vergelijk het Nederlandse categorale stelsel) in het voortgezet onderwijs
- De evaluatiemogelijkheden van het systeem (gedefinieerd als de hoeveelheid en intensiteit van bepaalde vormen van evaluatie, zoals nationale toetsingsprogramma’s, tentamens, schoolinspectie, een beheersysteem voor onderwijsinformatie, etc.)
- De omvang en diversificatie van de onderwijsondersteunende structuur in een land (mogelijk bestaande uit een eenheid die curricula ontwikkelt, ICT-services, schooladvies, een eenheid die het onderwijs evalueert en toetst, etc.)
- De verhouding van particulier onderwijs en door de overheid gesubsidieerd onderwijsbeleid waarbij stimulansen worden gebruikt om de prestaties van de school te verbeteren
- De mate waarin de schoolkeuze vrij is

Tabel 5.2: Indicatoren van financiële en materiële middelen op systeemniveau (input indicatoren)

Indicatoren van financiële en materiële “resources” op systeemniveau
<ul style="list-style-type: none"> ▪ Percentage van bbp (bruto binnenlands product) dat wordt besteed aan onderwijs ▪ Onderwijsuitgaven per leerling ▪ Verhouding publieke/particuliere investeringen in onderwijs ▪ Publieke investering in onderzoek en ontwikkeling op het gebied van onderwijs ▪ Totale uitgaven aan programma’s en speciale voorzieningen voor kansarme leerlingen ▪ Aanbod aan ondersteunende diensten door de overheid ▪ Particuliere bestedingen en overheidssteun aan ouders ▪ Percentage dat wordt besteed aan salarissen voor administratief personeel ▪ Percentage dat wordt besteed aan pensioenen voor onderwijzend personeel ▪ Percentage dat wordt besteed aan salarissen voor docenten

Tabel 5.3: Verschillende soorten maatschappelijke omstandigheden die relevant zijn voor onderwijs (context indicatoren)

Contextuele omstandigheden van onderwijssystemen
<ul style="list-style-type: none"> ▪ Demografische ontwikkelingen ▪ De arbeidsmarkt (bijvoorbeeld tekorten of overschotten in bepaalde sectoren) ▪ De algemene toestand van de economie ▪ Inkomensongelijkheid ▪ Relevante culturele aspecten ▪ De institutionele infrastructuur ▪ De volksgezondheid in een land ▪ Natuurrampen en oorlog

Sommige van de maatschappelijke dimensies uit bovenstaande tabel kunnen worden vertaald naar specifieke omstandigheden binnen het onderwijssysteem. In onderstaande tabel wordt een overzicht gegeven van aspecten die relevant zijn voor beschrijving en de ontwikkeling van indicatoren.

Tabel 5.4: Verdere uitwerking contextuele condities binnen het onderwijssysteem

Demografie
<ul style="list-style-type: none"> ▪ Vraag naar en aanbod van docenten in een land ▪ Het percentage docenten ouder dan 50 ▪ De verdeling naar geslacht van het docentencorps op schoolniveau ▪ Het percentage leerlingen op een school waarvan de leeftijd afwijkt van de indeling in groepen ▪ Heterogeniteit van de leerlingenpopulatie, culture herkomst, etniciteit
Culturele aspecten
<ul style="list-style-type: none"> ▪ De perceptie van de status van docenten door het grote publiek ▪ Waardering voor onderwijs en les krijgen ▪ Verwachtingen met betrekking tot de pedagogische functies van de school (bijvoorbeeld staatsburgerlijke vorming, ethische vorming, les over democratie) ▪ Cultureel bepaalde ideeën met betrekking tot autoriteit en leiderschap binnen het onderwijs
Institutionele infrastructuur
<ul style="list-style-type: none"> ▪ De mate van formalisering van de werkomstandigheden van docenten ▪ De formalisering van de rechten van docenten, leerlingen en ouders (bijvoorbeeld vrije schoolkeuze) ▪ Formele controle en inspectie op scholen ▪ Regels, en toezicht op de naleving daarvan, met betrekking tot verzuim van docenten ▪ Regelgeving met betrekking tot privéonderwijs door docenten uit het openbaar onderwijs ▪ Maatregelen voor fraudebestrijding in het onderwijs ▪ Kader voor vaststelling en evaluatie van het curriculum

Omdat de maatschappij en zelfs de nationale cultuur “als geheel” van invloed zijn op het onderwijssysteem is er sprake van een grote verscheidenheid aan variabelen, die allemaal behoren tot het domein van de context van het onderwijsstelsel. Voor de volledigheid voeg ik in de onderstaande tabel een overzicht van opbrengstindicatoren toe, eveneens ontleend aan Scheerens (2011).

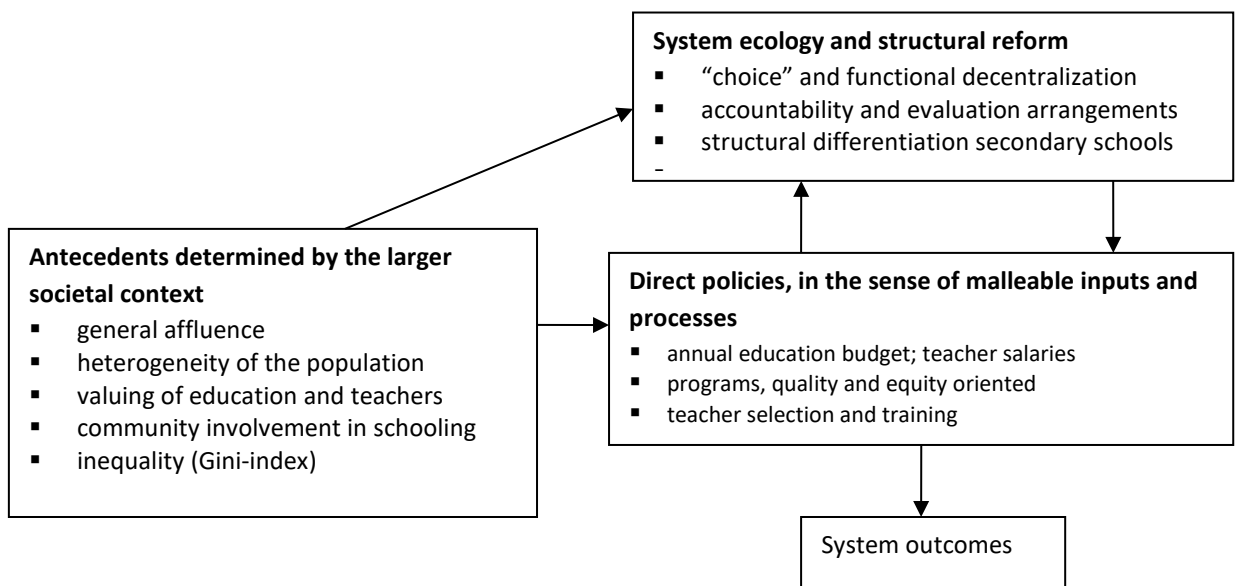
Tabel 5.5: Overzicht van onderwijs opbrengst-indicatoren

Hoofdcategorie opbrengst-indicator	Subcategorieën	Technische aspecten
Output-indicatoren	Prestatiemeting <ul style="list-style-type: none"> - inhoudelijk - lees-, reken- en denkvaardigheid - competenties (Bijvoorbeeld 'leren leren') 	<ul style="list-style-type: none"> - Meting toegevoegde waarde; groeicurven - Evaluatiemethodologie (variërend van multiple choice toets tot authentieke evaluatie) - Criteriumgericht dan wel normgericht toetsen
Outcome-/ attainment indicatoren	Meting opleidingsniveau <ul style="list-style-type: none"> - slagingspercentages - percentage geslaagde leerlingen zonder vertraging - aantal ongekwalificeerde schoolverlaters - percentage zittenblijvers 	Controlefactor: op selectie gericht schoolbeleid
Impact-indicatoren	Sociale participatiegraad <ul style="list-style-type: none"> - (voor elk scholingsniveau) percentage met betaald werk op bepaald niveau - percentage werklozen - (voor lagere opleidingsniveaus) percentage ingeschrevenen bij vervolgopleiding - sociale participatiegraad (sociaal kapitaal) - alfabetiseringsniveau - gemiddeld inkomen voor elk attainmentniveau; inkomstenverschillen - tekort/overschot vaardigheden 	Beschikbaarheid nationale onderwijs- en arbeidsmarkt cijfers metingen van sociaal kapitaal en alfabetiseringsniveau

In de volgende paragraaf gaan we verder in op een beperkter aantal kernvariabelen, wat betreft focus op opbrengsten, enige basis van empirische evidentie daarover en theoretische onderbouwing.

5.2.2 Een beperkter aantal kernfactoren

In de onderstaande figuur 5.1, ontleend aan Scheerens (2016, p 157), worden de voornaamste door het beleid te beïnvloeden variabelen (in overeenstemming met het systeemmodel in Fig. 2.1) verdeeld over twee categorieën: enerzijds systeem ecologische variabelen en anderzijds aandachtsgebieden van "lopend" onderwijsbeleid van de overheid. De eerste categorie bestaat uit structurele en geïnstitutionaliseerde kenmerken van onderwijssystemen, die weliswaar veranderbaar zijn, maar alleen door tijdrovende en ingrijpende hervormingen. De tweede categorie kan bij wijze van spreken op jaarbasis bijgesteld worden. In de figuur zijn alleen algemene labels van factoren opgenomen. In de bijbehorende tabel 5.6 zijn deze nader gespecificeerd.



Figuur 5.1: De systeem-context en beleidsfacetten

Tabel 5.6: *Systeemniveau, context, ecologie en beleidsfacetten. Naar, Scheerens, 2016, p.158*

Antecedents determined by the larger societal context	System ecology and structural reform	Direct educational policies in the sense of malleable inputs and processes
<p><u>General affluence of the country/region</u> GDP GDP per capita Annual unemployment quota (number of unemployed of the total potential working population) Percentage of people receiving social support of the potential adult working population)</p> <p><u>Heterogeneity of the population</u> Proportion of first, second and third generation immigrants Proportion of immigrants from developing countries</p> <p><u>Status of teachers</u> Attractiveness of the teaching profession relative to other professions, as rated by the general public and teachers</p> <p><u>Community involvement in schooling</u> Proportion of decisions that are taken by the local community/municipality (locus of decision making survey)</p> <p><u>Education relevant aspects of national cultures</u></p> <p><u>Societal (in)equality of country or region</u> (Gini index Measures the extent to which the distribution of income (or consumption) among individuals or households within a country deviates from a perfectly equal distribution.)</p>	<p><u>“Choice”</u> Free school choice Private/public schools</p> <p><u>Evaluation and accountability arrangements</u> Variety of methods Institutional Infrastructure High or low stakes nature of evaluation and accountability arrangements Whether or not a system has a standard based examination at secondary level</p> <p><u>Functional decentralization</u> (de) centralization in curriculum, primary process of teaching, human resources management and financial management</p> <p><u>Structural differentiation of secondary education</u></p> <p><i>Structural provisions</i></p> <ul style="list-style-type: none"> . age of first selection of secondary education . range of school types at secondary level (comprehensive versus categorical systems) . percentage of schools that use streaming, in the sense of ability grouping per classroom 	<p><u>Investment in education</u> Percentage of GDP spent on education Per pupil expenditure per ISCED level Teacher salaries. Annual education budget</p> <p><u>Teacher selection and training.</u> Mandatory level of training Professional development Entrance levels teacher training programs</p> <p><u>Quality oriented reform programs</u> Common core standard based curricula Evidence based approaches (Comprehensive School Reform) Bottom up reform</p> <p><u>Equity oriented policies</u> <i>Compensatory measures for disadvantaged learners</i></p> <ul style="list-style-type: none"> . proportion of education budget spent on special programs . priority of stratification of specific disadvantaged groups . preferred strategies to compensate for educational inequality

5.3 Conceptualisering en theoretische onderbouwing van kernfactoren

Bij de beschrijving van kernfactoren, die zijn opgenomen in Fig. 5.1, richt ik mij op structurele kenmerken en facetten van lopend beleid. Wat de structurele kenmerken betreft gaat het om basisconstellaties van het onderwijsbestel, zoals de mate van centralisatie en decentralisatie, evaluatie en accountability voorzieningen en de opbouw van het scholenstelsel. Met “facetten van lopend beleid” wordt bedoeld op maatregelen op de korte en middellange termijn gericht op beleidsterreinen als de jaarlijkse onderwijsbegroting, de initiële lerarenopleiding, nationale programma’s om de kwaliteit en “gelijkheid” van het onderwijs te stimuleren en initiatieven tot vernieuwing van het curriculum. Voor een nadere toelichting bij de categorie van gegeven condities die door de ruimere maatschappelijke omgeving wordt bepaald – de linker kolom uit tabel 5.6 - zij verwezen naar Scheerens, (2016, p. 148 e.v.).

5.3.1. Inleiding, verschillende interpretaties van het rationaliteitsparadigma

Besturings- en planningsmodellen kunnen worden gezien als varianten van het rationaliteitsparadigma. Zuivere rationaliteit wordt gekenmerkt door volledige kennis over operationele doelen en optimale methoden om deze doelen te bereiken. Dit ideaal wordt zo dicht mogelijk benaderd in het concept van proactieve synoptische planning. Kenmerken zijn van tevoren vastgestelde doelstellingen, methoden die onderbouwd zijn door wetenschappelijk onderzoek, gecontroleerde implementatie en onafhankelijke evaluatie. Praktische voorbeelden in het onderwijs zijn centralistische onderwijsstelsels met een gedetailleerd voorschrijvend curriculum, het idee van “reforms as experiments” (Campbell, 1969) en op schoolniveau het model van “High reliability organizations” beschreven in hoofdstuk 4. Op het niveau van de centrale overheid staat dit model ver van de realiteit van het Nederlandse onderwijs en daarom zal er hier geen verdere aandacht aan worden besteed. Wij gaan daarom in op varianten van het rationaliteitsparadigma, waarin alternatieve mechanismen worden gebruikt om te komen tot effectieve doelbereiking. Een overzicht wordt gegeven in Tabel 5.7.

Tabel 5.7: Interpretations of the rationality paradigm

Interpretation of the rationality paradigm	mechanism
Synoptic planning	Proactive structuring
Cybernetics	Evaluation and feedback
Public choice theory	Market mechanisms; school competition, choice

5.3.2. Keuze-vrijheid en functionele decentralisatie⁸

Positieve verwachtingen over de effectiviteit van decentralisatie komen min of meer overeen met soortgelijke verwachtingen over de voordelen van geprivatiseerd onderwijs. In beide gevallen wordt ervan uitgegaan dat autonomere scholen responsiever zijn ten opzichte van de verwachtingen van de “cliënten” van het onderwijs en de lokale context en daarmee minder bureaucratisch en meer resultaatgericht. In veel geïndustrialiseerde landen heeft gedurende de laatste 3 decennia decentralisatie in het onderwijs plaatsgevonden, al zijn er ook landen die een golfbeweging van centralisatie en decentralisatie laten zien (bijvoorbeeld Zweden en Engeland). Het Nederlandse

⁸ De term decentralisatie wordt hier als een overkoepelende term gebruikt voor het verplaatsen van centrale invloed naar lagere niveaus. Fijnere specificaties in de zin van deconcentratie, delegatie en deregulering worden hier niet toegelicht, vgl. Scheerens, 2016, p 150

onderwijs heeft, met een beleid dat historisch gezien sterk werd medebepaald door de schoolstrijd en principes als “soevereiniteit in eigen kring”, een traditie van autonome scholen. Wel is daar in de jaren negentig van de vorige eeuw verdere deregulering en autonomievergroting in het financiële domein aan toegevoegd.

De theoretische onderbouwing van de effectiviteit bevorderende werking van decentralisatie en autonomie wordt gegeven in de politieke visie op het functioneren van organisaties (Morgan, 1986) en de economische “public choice” theorie (Niskanen, 1971). In de theorie van Morgan wordt de veronderstelde gemeenschappelijke gerichtheid op de organisatie doelstellingen geproblematiseerd. Het feit dat in deze organisaties sterke incentives en controles ontbreken die gemeenschappelijkheid afdwingen betekent dat er ruimte is om eigen belangen na te streven. Ook in de theorie van Niskanen staat het probleem van effectieve controle door democratisch gekozen besturen centraal. Het creëren van competitie en marktwerking worden hier als remedie gezien. In het onderwijs wordt dit bewerkstelligd door het instellen van privéscholen en vrije schoolkeuze. Het veronderstelde voordeel van autonome scholen die marktgericht opereren is verder dat het ontbreken van overheidsregulering creativiteit en responsiviteit stimuleert (Chubb & Moe, 1990).

In het internationaal vergelijkende onderzoek ontstond de behoefte aan een nuancering van de tweedeling centraal-decentraal.

In het kader van het OECD- indicatorenproject werd nadere uitwerking gegeven aan het concept “functionele decentralisatie”. Dit concept specificiert de mate van centralisatie/decentralisatie als een continuüm, dat apart voor functionele gebieden binnen het onderwijsbeleid kan worden gespecificeerd (Van Amersfoort en Scheerens, 1997). In het INES project van de OECD werd de “locus of decision making” vragenlijst ontwikkeld om functionele decentralisatie te meten als internationaal vergelijkende indicator. Met dit instrument werden vier functionele domeinen van onderwijsbeleid onderscheiden: onderwijs, formele regulering, personeelsmanagement en financieel management. Per besturingsniveau wordt nagegaan welke beslissingsbevoegdheid er is op elk van de domeinen. In het lopende PISA-onderzoek wordt nog steeds van deze indeling gebruik gemaakt, zij het in aangepaste vorm. De onderwijsdimensie heeft vooral betrekking op de speelruimte van leerkrachten bij het onderwijs. Het domein van de formele regulering, waarin bijvoorbeeld de vraag naar het centrale karakter van het curriculum past, wordt niet meer als zodanig gebruikt, al past de vraag of onderwijsstelsels een op standaarden gebaseerd centraal examen hebben hier wel bij. De categorieën personeelsmanagement en financieel beheer komen nog wel voor. In hoofdstuk 3 is ingegaan op curriculum alignment. In de VS heeft dit begrip een zekere opleving meegemaakt met de introductie van het “common core” nationale curriculum (Vgl. Polikov en Porter, 2016). Zoals aangegeven in Hoofdstuk 3 leent het zich voor een systemische uitwerking, waarbij onder meer interessant is welke invloed de centrale overheid heeft over nationale doelstellingen en standaarden en in welke mate er regulering bestaat bij het aanbod van leerboeken, methoden en onderwijstechnologie. Eindtoetsen en examens hebben een centrale plaats bij curriculum alignment.

Privatisering van het onderwijs wordt vaak geoperationaliseerd als het aandeel privéscholen in vergelijking met het aantal “publieke” of openbare scholen. Echter in het kader van het streven naar marktwerking in het onderwijs zou het ook interessant zijn om de invloed van private ondernemingen op de productie van leermiddelen na te gaan.

Een bijzondere Nederlandse vorm van delegatie van beslissingsbevoegdheid is de functie van de Sector Raden, die formeel de functie van werkgever in het onderwijs hebben gekregen.

5.3.3 Evaluatie en Accountability

Het “vrijmaken” van processen en het behouden van controle over opbrengsten is te zien als een patroon van functionele decentralisatie dat de bestuurlijke kant van evaluatie en het “rekenschap doen geven” (accountability) benadrukt. Dit patroon van functionele organisatie komt overeen met het besturingsmodel dat bekend staat als “new public management”. Het meten van opbrengsten van onderwijs en het gebruiken van de informatie die dit oplevert kan op drie verschillende manieren worden gezien als een stimulans van kwaliteit en effectiviteit. De klassieke betekenis van “accountability” benadrukt administratieve controle als mogelijkheid om de kwaliteit te borgen en bij te sturen. Glass (1972) stelde vast dat accountability bestaat uit drie constituerende deelprocessen: het openbaar maken van de opbrengsten van dienstverlening, het toetsen en beoordelen van de geleverde prestatie en wat hij aanduidt als “redress for poor performance”. Dit laatste element verwijst naar sancties, zoals bijvoorbeeld het sluiten van een school na herhaald ondermaats presteren.

Een tweede betekenis van accountability past bij de economische “public choice” theorie en het principe van de marktwerking. Hierbij wordt aangenomen dat vrije schoolkeuze geleid wordt door informatie over de prestaties van een school. De markt zou scholen op die manier dwingen om hun prestaties op peil te houden.

Analyses en evaluaties van beide invullingen van “accountability” wijzen uit dat beide benaderingen in de praktijk minder effectief zijn door implementatieproblemen. Volgens Cibulka en Derlin (1995) vindt de administratieve variant weinig daadwerkelijke toepassing. Met name de “sanctie”-component wordt niet graag gehanteerd. Bij de marktwerkingsvariant blijkt vaak een belangrijke voorwaarde niet vervuld, namelijk de aanname dat informatie over opbrengstkwaliteit bepalend is bij de schoolkeuze (vgl. Scheerens, 2011).

Volgens Bosker en Scheerens (1999) is er een derde interpretatie die de kwaliteitbevorderende werking van accountability zou kunnen verklaren. Deze is gebaseerd op de aanname dat, ook bij het ontbreken van sterke externe incentives (zoals een straffende overheid, en kritische ouders) scholen gevoelig zijn voor openbare publicatie van resultaten. Voorbeelden zijn de vroegere kwaliteitskaarten van de onderwijsinspectie en schoolranglijsten in publieke media.

Naast deze drie varianten van “accountability” zijn er nog enkele andere interpretaties, die soms als accountability worden geafficheerd, maar beter te zien zijn als specifieke toepassingen van het gebruik van evaluatieve informatie. Zo bespreken Ellmore et al. (1990) prestatiemeting en feedback als onderdeel van professionele ontwikkeling en hanteren daarvoor de term “professional accountability”. Dit is een perspectief dat eerder geassocieerd is met organisatie leren en feedback dan met “ter verantwoording roepen” door externe partijen. Het gaat hier om een vorm van interne, in tegenstelling tot externe evaluatie, te vergelijken met interne kwaliteitszorg en formatieve school zelfevaluatie. Carnoy et al. (2003) spreken in deze van “interne accountability”; waarbij ze vaststellen dat deze gestimuleerd kan worden door “harde” externe accountability.

Behalve als onderdeel van accountability en intern organisatie-leren heeft opbrengstevaluatie een belangrijke, formeel geïnstitutionaliseerde functie namelijk in de vorm van certificatie. Centrale examens zijn de bekendste uitwerkingen hiervan. Examens en diplomering geven rechten aan leerlingen en maken duidelijk wat de maatschappij mag verwachten van het onderwijs (Zie ook Onderwijsraad, 2005).

Tot slot een resumé van de mechanismen die maken dat centrale examens en “new public management” gezien kunnen worden als hefbomen van beleid dat op handhaving en verbetering van kwaliteit is gericht:

- Examens reguleren het civiel effect van onderwijs.
- Examens stimuleren de (extrinsieke) motivatie van leerlingen.
- De resultaten van examens kunnen gebruikt worden voor de externe beoordeling van scholen.
- De resultaten van examens kunnen tevens gebruikt worden voor interne toepassing door scholen als feedback op het eigen functioneren.
- Examens bieden een curriculaire focus voor het onderwijs.

Het besturingsmodel van “new public management” is aantrekkelijk voor onderwijsstelsels waar sprake is van een grote mate van vrijheid bij de inrichting van processen. Enerzijds wordt er veel speelruimte geboden, anderzijds is er controle op de dimensie waar het uiteindelijk om gaat in het onderwijs: de optimalisering van leerprestaties. Opbrengstgerichte of evaluatiegecentreerde sturing vormt een middenweg tussen proactieve, synoptische planning enerzijds en het marktgerichte “public choice model”.

5.3.4 De stratificatie van scholenstelsels; geïntegreerd versus categoriaal

Dit kenmerk van scholenstelsel komt niet direct voort uit een bepaald theoretisch principe, maar heeft wel de aandacht getrokken in internationaal vergelijkende assessment onderzoeken, met name PISA. Het blijkt namelijk dat geïntegreerde scholenstelsels in het voortgezet onderwijs vaak beter scoren op gelijkheidsindicatoren dan sterk gestratificeerde scholenstelsel. Soms blijken ze het zelfs niet alleen beter te doen op gelijkheid, maar ook op kwaliteit (cq hogere leerprestaties). Wel is het interessant om zich af te vragen hoe deze uitkomst te verklaren is. Onderzoek dat binnen scholen is uitgevoerd naar het groeperen van leerlingen, enerzijds in heterogene groepen en anderzijds in meer homogene op niveau van presteren ingedeelde groepen, laat zien dat heterogeen groeperen over het algemeen beter werkt voor zwakkere leerlingen. Dit wordt op zijn beurt weer verklaard door de veronderstelling dat in heterogene groepen de norm voor presteren hoog wordt gehouden (De Vos, 1989). Het lager presteren, met name aan de onderkant van de verdeling, in sterk gestratificeerde stelsels wordt ook wel toegeschreven aan het “opsluiten” van leerlingen in aparte scholen, zeker als dit op relatief jonge leeftijd gebeurt (Luyten et al. 2005). Als indicatoren van stratificatie worden daarom meestal zowel het aantal schoolcategorieën gebruikt als de leeftijd, waarop de keuze moet worden gemaakt voor een bepaalde categorie voortgezet onderwijs (vgl Van Ravens, 2011). Het thema van de vroege schoolkeuze en sterke stratificatie is te zien als een zich chronisch voortslepende kwestie in het Nederlandse onderwijs. Na het mislukken van het Middenschool- experiment in de jaren zeventig en de daaropvolgende Basisvorming lijkt de keuze voor drastische oplossingen uitgesloten. De implementatie van incrementele oplossingen als een verlengde brugperiode verloopt al jaren moeizaam, maar blijft terecht een aandachtspunt

5.3.5 Facetten van “lopend” beleid

Met lopend beleid wordt niet specifiek bedoeld op beleid dat momenteel nog in uitvoering is, maar op maatregelen en interventies die mogelijk zijn als onderdeel van het permanente beleidsarsenaal van de overheid zonder dat daarvoor stelselherzieningen of wijzigingen van centrale onderwijsdoelstellingen nodig zijn (vergelijk de factoren die worden genoemd in de derde kolom van Figuur 5.6). Daarbij is het soms enigszins arbitrair om beleid enerzijds te classificeren als structureel en alleen te veranderen door ingrijpende structuurhervormingen en anderzijds als meer fluctuerend en op kortere termijn te wijzigen.

Wanneer we bijvoorbeeld in deze subparagraaf de initiële opleiding van leraren aan de orde stellen, zijn beide interpretaties mogelijk. Wanneer men zou besluiten om alle lerarenopleidingen op universitair niveau te brengen, dan zou dat eerder te zien zijn als een structurele hervorming, dan wanneer er bijvoorbeeld besloten zou worden tot een bijstelling van het curriculum. Ook curriculumvernieuwing kan zowel het karakter hebben van een structurele hervorming als van een meer incrementele bijstelling. Men zou van curriculumvernieuwing kunnen spreken als er zowel sprake is van verandering van middelen als van doelen en van curriculumverbetering wanneer de doelen onveranderd blijven maar de middelen geoptimaliseerd worden.

Er zal aandacht worden besteed aan de volgende facetten van lopend beleid:

- De onderwijsbegroting.
- Nationale beleidsprogramma's die gericht zijn op kwaliteitsverbetering.
- Initiatieven tot curriculumvernieuwing.
- Beleid gericht op het tegengaan van achterstanden en ongelijkheid in het onderwijs.
- De initiële opleiding van leerkrachten.

Om verschillende redenen is de behandeling van facetten van lopend beleid minder uitvoerig dan de voorbeelden van structurele factoren in de vorige paragraaf. Ook beperk ik mij in de volgende paragraaf over empirische evidentie tot indrukken op hoofdlijnen. De kern van deze studie heeft betrekking op de resultaten van empirisch onderwijseffectiviteitsonderzoek en de theoretische onderbouwing van die resultaten. Als het gaat om de bovengenoemde facetten van "dagelijks" beleid zou de empirische evidentie gebaseerd moeten zijn op beleidsevaluaties (die soms wel maar soms ook niet zijn uitgevoerd). Dit is zeker een aanpalend gebied, maar binnen de beperkingen van deze studie is een diepgaande beschrijving niet uitvoerbaar (gedeeltelijk is dit wel beoogd in Scheerens, 2016, hfst 12). Dit betekent dat concluderende uitspraken veelal als hypotheses moeten worden opgevat.

De onderwijsbegroting

In het kader van internationaal vergelijkend onderzoek worden de onderwijsuitgaven van landen vaak uitgedrukt als het percentage van het bruto nationaal product dat besteed wordt aan onderwijs. Fluctuaties hiervan kunnen vergeleken worden met eventuele veranderingen in onderwijsopbrengsten, waarbij de aanname is dat de productiviteit toeneemt naarmate het budget hoger is. Relevante uitsplitsingen van de uitgaven, zoals een verdeling tussen uitgaven aan het primaire onderwijsproces enerzijds en overhead in de zin van kosten voor bestuur, management en ondersteuning, zijn lastig internationaal te vergelijken.

Op kwaliteitsverbetering gerichte nationale programma's

Voorbeelden van dit soort programma's gedurende de laatste twee decennia zijn:

- De zogenoemde kwaliteitsagenda's van de overheid in samenwerking met de sectorraden (Vgl. OCW, 2007 en 2008).
- Het programma "School aan Zet".
- Het programma "Zicht op Kwaliteit".
- Initiatieven om "opbrengstgericht werken" te stimuleren.

Het is buiten de reikwijdte van deze studie om deze initiatieven nader te bespreken. In de volgende paragraaf, waarin gekeken wordt naar empirische evidentie zal wel een overkoepelende karakterisering worden gegeven

Initiatieven tot curriculumvernieuwing

Op dit terrein zijn te noemen het ontwerp “Onderwijs 2032” en het nog lopende “curriculum.nu”. De outline die in “Onderwijs 2032” wordt gegeven moet worden opgevat als een structurele hervorming; in hoeverre dat ook het geval is bij “curriculum.nu” is nog niet zo duidelijk.

Beleid gericht op het tegengaan van onderwijsachterstanden

Beleid dat gericht is op het terugdringen van onderwijsachterstand heeft een lange traditie in Nederland. De kern van het beleid bestaat uit het verstrekken van extra financiële middelen aan scholen met een hoog percentage achterstandsl leerlingen. Door de jaren heen zijn de formules voor deze extra toewijzing verschillende keren gewijzigd. Scholen zijn vrij in de besteding van deze middelen. In de praktijk worden ze voornamelijk gebruikt voor klassenverkleining. De laatste decennia is het terugdringen van achterstanden in het kader van de voor- en vroegschoolse opleiding een nieuwe pijler van het beleid geworden.

De initiële opleiding van leerkrachten

Het belang van goed opgeleide leerkrachten hoeft geen betoogd. In hoofdstuk 2 is het belang van “de factor leerkracht” ook nader kwantitatief onderbouwd op basis van economisch onderzoek van onder meer Hanushek en Rivkin. Datzelfde hoofdstuk biedt ook een empirisch en theoretisch onderbouwd beeld van effectieve instructie; waarbij het accent ligt op het gedoseerd aanbieden van structuur en ondersteuning. Het zou wenselijk zijn dat deze kennis over effectieve instructie wordt overgebracht in de lerarenopleiding. Tegelijkertijd zagen we in Hoofdstuk 3 dat in populaire visies op de school als lerende organisatie veel nadruk wordt gelegd op allerlei secundaire functies van leerkrachten. Het zou interessant zijn om te weten hoe de verhouding tussen aandacht voor het primaire onderwijsproces en aandacht voor de meer secundaire functies tot uitdrukking komt in het curriculum van de lerarenopleiding. Entree-vaardigheden van instromers en het niveau van de opleiding zijn eveneens hefboomen om te sturen op de kwaliteit van de lerarenopleiding.

5.4 Resultaten van empirisch onderzoek

Over onderzoek naar systeemeffectiviteit bestaan, voor zover mij bekend, nauwelijks meta-analyses. De belangrijkste kennisbron zijn daarom op zichzelf staande onderzoeken. Een belangrijke deel bestaat uit secundaire analyses op internationale databestanden, naast directe resultaten die vermeld staan in de basisrapporten van deze studies. In dit hoofdstuk worden resultaten uit PISA gebruikt. In de Annex bij dit rapport van Meelissen, worden resultaten van TIMSS gepresenteerd. Als “advance organizer” van de inhoud van deze paragraaf wordt hieronder een overzicht weergegeven dat ontleend is aan Scheerens, Luyten en Van Ravens, (2011).

Autonomie van scholen Wisselende uitkomsten, meestal niet significant (Scheerens en Maslowski, 2008). Openbaar versus privéscholen. Geen effect, nadat gecontroleerd is voor de samenstelling van de leerlingenpopulatie (Luyten en anderen, 2005)

Accountability en een goed ontwikkeld examensysteem Vergelijk het werk van Bishop en Woessmann, Scheerens et al., 2009 geven een overzicht. Overwegend positieve impact; discussie over neveneffecten van accountability.

Stratificatie Categorale versus geïntegreerde (*comprehensive*) schoolstelsels in het vo. Overwegend negatieve effecten, dat wil zeggen dat sterk gestratificeerde stelsel minder gelijkheid laten zien, en soms ook lagere scores, zie bijvoorbeeld Brunello en Checchi, 2006, Jakubowski, 2009, Causa & Chapuis, 2019.

Figuur 5.2 Overzicht ontleend aan Scheerens, Luyten en Van Ravens, (2011)

In het vervolg wordt een uitvoeriger bespreking van onderzoeksresultaten gegeven, die berust op Scheerens, 2016, hfst 9.

5.4.1 **Illustratieve onderzoeksresultaten**

Achtereenvolgens worden resultaten samengevat over de effecten van school autonomie, accountability, vrije schoolkeuze, aandeel private scholen in een land en de stratificatie van scholenstelsels in het voortgezet onderwijs.

5.4.1.1 *Keuzevrijheid, privé-scholen en autonomie*

Meer vrijheid en beslissingsbevoegdheid van scholen als resultaat van overheidsbeleid dat decentralisering vergroot, heeft verschillende verschijningsvormen. Een eerste vorm is de vergroting van het aantal privé-scholen. De twee andere vormen kunnen ook worden toegepast binnen het publieke onderwijs. Vrije schoolkeuze kapitaliseert op de mogelijkheid dat het marktmechanisme in werking wordt gesteld om de effectiviteit van scholen te vergroten (zie de eerdere behandeling van de publieke keuze theorie). De derde mogelijkheid gaat uit van meer beslissingsbevoegdheid van de schoolleiding. Hieronder worden resultaten van empirisch onderzoek vermeld, voor elk van deze drie vormen.

Vrije schoolkeuze

Een baanbrekend onderzoek naar de veronderstelde voordelen van privéonderwijs was de studie van Chubb & Moe (1990). Hun onderzoek wees uit dat in privéscholen hogere besturingsniveaus, waaronder het schoolbestuur en toezichthouders, minder beslissingsbevoegdheid hebben dan in het publieke onderwijs. Ook bleken privéscholen meer speelruimte bij het personeelsbeleid te hebben. Het was eenvoudiger om nieuw personeel in te huren dan om zittend personeel te ontslaan. Zij stelden verder vast dat schoolleiders in privéscholen vaker een onderwijsachtergrond en lesgeef-ervaring hadden dan schoolleiders in openbare scholen. Ook manifesteerden schoolleiders in privéscholen vaker onderwijskundig leiderschap. Op het onderzoek van Chubb & Moe volgde forse methodologische kritiek. (Glass & Matthews, 1991; Smith & Meier, 1995; en Witte, 1990). Smith & Meier legden het onderzoek van Chubb & Moe onder de loep en onderzochten de effecten van programma's voor vrije schoolkeuze in andere landen. Zij kwamen onder andere tot de conclusie dat het pleidooi van Chubb & Moe voor het gebruik van vouchers niet voldoende empirisch onderbouwd is. Deze kritische conclusies werden verder

gesteund op basis van meta-analyses over privéonderwijs en vouchers in de Verenigde Staten en Zuid-Amerika (McEwan, 2000), Somers, McEwan & Willms, 2004).

Privéscholen

In de hierboven aangehaalde onderzoeken van McEwan en Somers et al. bleek dat in de vergelijking tussen openbare en privéscholen, de samenstelling van de leerlingenpopulatie een grote invloed had. Privéscholen trekken doorgaans een qua thuismilieu meer bevoorrechte groep. Wanneer hiervoor in de statistische analyses wordt gecontroleerd, verdwijnen de verschillen in prestaties. In het onderstaande citaat, ontleend aan een internationale studie van Luyten et al. (2005) worden naast openbare en privéscholen ook zogenoemde “*government dependent private schools*” betrokken.

“Levels of autonomy differ between private, government dependent private and public schools. Independent private schools enjoy more autonomy in all domains of decision making (personal management, learning and instruction, planning and structures and human and financial resources). According to the decision domain, government-dependent private schools fall between independent private and public schools. With regard to approving students for admittance to school, formulating the school budget and deciding on budget allocations, as well as the appointment and dismissal of teachers, government-dependent schools are very much like independent private schools. Concerning decisions on teachers’ starting salaries and salary increase, the determination of the course content and student disciplinary policies, government-dependent private schools show great resemblance to public schools. Compared to other school types, public schools are less favorably endowed with educational and physical resources, employ teachers with a (comparatively) low morale, and are characterized by poorer school climates. Independent private schools report the most favorable resources and climate, followed by government-dependent private schools. In most countries teaching and learning in public schools takes place under less advantageous conditions than in private schools. These differences in school conditions imply differential educational opportunities for students attending different school types. However, despite these disadvantages, in half of the participating countries, there are no significant performance differences between students in public schools as compared to students in independent private schools. For the remaining countries students in independent private schools outperform students from public schools. The school composition again plays the most significant role in these performance advantages: independent private schools lose their performance advantage in all countries, once student and school characteristic are taken into account. This is confirmed by the few countries in which public schools significantly outperform independent private schools: controlling for student and school characteristics show that this is largely attributable to a more favorable intake” (ibid, p. 84).

Dit citaat laat zien dat er wel degelijk verschillen zijn in voorzieningen tussen openbare scholen en privéscholen, maar dat deze verschillen niet aantoonbaar doorwerken in de leeropbrengsten.

Autonomie als vergrote beslissingsbevoegdheid van scholen

Ik bespreek een drietal onderzoeklijnen die betrekking hebben op de effectiviteit van autonomievergroting.

- a) *Onderzoek gericht op het effect van vergrote autonomie van scholen op intermediaire school variabelen, die op hun beurt te beschouwen zijn als effectiviteit bevorderend*

Winkler and Gershberg (2000) stelden vast dat decentralisatie resulteerde in hoger gekwalificeerde schoolleiders, die sterk gericht waren op schoolverbetering. Levačić (1995) en Whitty, Power and Halpin (1998) gebruikten een soortgelijke onderzoekszopzet, maar kwamen tot andere conclusies. Deze auteurs stelden namelijk vast dat toegenomen autonomie niet leidde tot meer focus op schoolverbetering, door middel van meer aandacht voor samenwerking van stafleden.

- b) *Evaluatieonderzoek naar “school based management” en systeem-brede vernieuwing*
Onderzoek van Malen, Ogawa & Kranz (1990) en van Summers and Johnson (1996) leverde geen bewijs dat “school-based management” effectief is in het bevorderen van leerprestaties. Review studies van Leithwood en Menzies (1998a, b), Fullan en Watson (2000) en Bryk, Thum, Easton & Luppescu, (1998) leverden geen eenduidige resultaten op voor het effect van grotere autonomie maar wezen wel op het belang van interactie-effecten, in de zin van andere variabelen die het effect van autonomievergroting leken te stimuleren. In het geval van het onderzoek van Bryk et al. betrof dit met name rigoreuze externe accountability-eisen.
Longitudinaal onderzoek naar systeem brede vernieuwingen in El Salvador en Nicaragua van Jimenez en Sawada (1998) constateerde slechts partiele effecten van decentralisatie en vergroting van schoolautonomie.
- c) *Onderzoek gebaseerd op internationaal vergelijken assessment onderzoek*
Walberg e.a. (2000) analyseerden de effecten van op decentralisatie gericht beleid in 14 landen. Ondanks overwegend negatieve conclusies vonden zij een significant en positief effect van autonomie op het terrein van het rekruteren van staffleden. Op basis van een secundaire analyse van data uit TIMSS 1995 data gebaseerd op OECD's (1998) “Locus of decision-making”-survey concludeerde Wößmann (2003) dat vergrote schoolautonomie een positief effect had op leerprestaties, maar alleen in onderwijsstelsels die een centraal examen in het voortgezet onderwijs hadden. Fuchs and Wößmann (2004), bevestigden deze uitkomst, dit maal op basis van PISA data. In hun onderzoek naar schoolfactoren geassocieerd met kwaliteit en gelijkheid rapporteerden Luyten et al. (2005) dat scholen met meer autonomie op het vlak van het personeelsbeleid beter presteerden. Dit effect verdween echter na controle voor de SES-gerelateerde samenstelling van de leerling-populatie. Alles bij elkaar zijn de uitkomsten van onderzoek naar de effectiviteit van verschillende facetten van schoolautonomie niet erg consistent. De uitkomst dat autonomie effectiever is, wanneer zij voorkomt in combinatie met externe accountability en examens, is bijzonder interessant en komt overeen met de principes van “new public management”. Deze uitkomst wordt mede voor een deel bevestigd in analyses op de data sets van PISA 2009, 2012 en 2015 (OECD, 2010, 2013, 2016). Zie ook de sub-paragraaf over accountability en de nadere bespreking van de desbetreffende PISA-rapporten verderop.

5.4.1.2 *Accountability*

Op basis van secundaire analyses op TIMSS-data liet Bishop (1997) zien dat landen met een op standaarden gericht examen gemiddeld beter presteerden van scholen zonder zo'n examen. Wößmann (2000), en Fuchs and Wößmann (2004) bevestigen het effect van examens. Rand News Release (2000) schreef vooruitgang in prestaties bij Amerikaanse Staten toe aan de intensiteit van de voorzieningen om rekenschap te geven van leerprestaties. Carnoy, Elmore en Siskin (2003) komen tot dezelfde conclusie. Wößmann et al., (2009) vonden positieve effecten van verschillende soorten evaluatieprocedures. Zo rapporteerden ze positieve effecten voor het bestaan van externe eindexamens, assessment in het kader van bevordering naar een hogere klas, assessment in het kader van vergelijking tussen scholen, de inspectie van lessen door de schoolleider en door externe inspecteurs. Daarnaast constateerden ze positieve effecten van accountability en evaluatie op gelijkheidsindicatoren (ibid, p. 84).
Ook Bol, Witsche, Van de Werfhorst en Dronkers (2014) laten zien dat de aanwezigheid van een centraal examen een matigend effect heeft op ongelijkheid in sterk gestratificeerde onderwijsstelsels.

Het positieve effect van centrale examens wordt opnieuw bevestigd in recentere overzichtsstudies van Woessmann (Woessmann, 2016, 2018). Overigens vinden niet alle onderzoeken positieve effecten van

examens en accountability (Scheerens et al. 2015). Verder zijn er onderwijsstelsels aan te wijzen die geen centraal examen hebben en toch zeer goede prestaties laten zien. Vlaanderen is een voorbeeld. Ook in schooleffectiviteitsonderzoek zijn de effecten van monitoring en evaluatie binnen scholen vaak onderzocht (Scheerens & Bosker (1997), Willms & Somers (2001), Scheerens et al., (2007). Zie ook de beschrijving hiervan in Hoofdstuk 4 van dit rapport. Andere auteurs wijzen op mogelijke neveneffecten van een hard accountability regiem (vergelijk Sacks, 1999; Cibulka & Derlin, 1995). Ook Honingh en Ehren (2012, p.69) refereren aan onderzoeken, waarin de effecten van strategisch gedrag bij toetsing in het onderwijs zijn onderzocht: “Het simplificeren van prestaties en hanteren van eenzijdige meetbare uitkomsten kan zelfs leiden tot onwenselijke gevolgen zoals ritualisme, bureaucratisering en een eenvormig beeld van kwaliteit, zo blijkt uit onderzoek naar (bij)effecten van prestatiemeting”.

Een nadere analyse van de achtergrond van deze verschijnselen wordt gegeven door Koretz, (2005) en Holcomb, Jennings en Koretz, 2013). In de Amerikaanse context is er sprake van “high stakes testing” in het kader van het programma, No Child Left Behind. Koretz constateert dat er sprake is van “test inflation” dat wil zeggen kunstmatig hoge toetsuitkomsten op lokale toetsen waarop scholen beoordeeld worden. Hij constateert dit door resultaten van dit soort ‘high stakes tests’ te vergelijken met resultaten op landelijke peilingen, waaraan voor scholen niet zulke grote belangen verbonden zijn. “Test inflation”, dat wil zeggen kunstmatig opgeschroefde toetsresultaten, ontstaat door gebrekkige inhoudsvaliditeit van toetsen, door strategisch gedrag van scholen en door “teaching to the test”. Verdere discussie over deze thematiek is te vinden in Scheerens en Exalto (2017) en Scheerens et al. (2019).

5.4.1.3 Stratificatie (categorale versus geïntegreerde scholenstelsels in het voortgezet onderwijs
Onderzoek naar het effect van stratificatie laat meestal negatieve uitkomsten zien, vergelijk Brunello & Checchi, (2006), Luyten et al. (2005), Causa & Chapuis, (2009). Geïntegreerde systemen doen het doorgaans beter dan sterk gestratificeerde, categorale stelsels. De voordelen blijken zowel als positieve effecten op het prestatieniveau, als op (on) gelijkheid. Als ongelijkheidsindicator wordt hierbij vaak de mate waarin leerprestaties bepaald worden door de sociaaleconomische status van de leerlingen gebruikt. Sommige onderzoeken vinden alleen effecten op ongelijkheid (vgl de resultaten van PISA 2009 en 2012) terwijl andere onderzoeken ook effecten op prestatieniveaus vinden (bv Luyten et al. 2005). Deze auteurs vatten de resultaten van hun analyses op PISA 2000 data als volgt samen:

“PISA 2000 also reveals that comprehensive education systems are not always more equitable in terms of variation in performance in reading literacy among students. There was no statistically significant differences between the average standard deviation in student performance of the countries with comprehensive education systems and the countries with the highest degree of institutional differentiation. However, countries with differentiated education systems are less equitable in terms of the impact of student socio-economic background on performance. Perhaps surprisingly, the PISA 2000 findings show more solid evidence for integrated, comprehensive school systems being high performers rather than champions of equity” (ibid, p.62).

Overigens doet Nederland het, als een van de sterkst gestratificeerde stelsels ter wereld, het ondanks deze “handicap” internationaal gezien relatief goed op “equity” (vgl Scheerens, Van der Werf en Timmermans, (2019). Overeenkomstige conclusies worden getrokken in een doorwrochte analyse van het Nederlandse scholenstelsel van de OECD, in het rapport over PISA 2015 (Vol. II, p. 173-175) onder het motto “Stratification policies in the Netherlands: context matters” (OECD, 2015. Van de Werfhorst en anderen (2015) concluderen dat, ook gegeven de sterke stratificatie, het feit dat Nederland een stevig centraal examen heeft een gunstige uitwerking heeft op de gelijkheid.

5.4.2 Aanvullende resultaten op basis van recente PISA-onderzoeken (2009, 2012, 2015 en 2018)

In deze paragraaf worden resultaten van PISA-onderzoek samengevat, die betrekking hebben op de systeemkenmerken die in het voorafgaande onderscheiden zijn; zie met name Tabel 5.4 de tweede kolom. De samenvatting is gebaseerd op een wat uitvoeriger beschrijving van gegevens uit 4 lichteningen van PISA, 2009, 2012, 2015 en 2018), zoals gepresenteerd in de desbetreffende rapporten van de OECD (zie **Annex 1**).

In de tweede sub paragraaf worden twee secundaire analyses besproken die gebaseerd zijn op gegevens uit PISA 2009. De bedoeling hiervan is om iets meer profiel te geven aan de betekenis van resultaten van dit soort studies, wat betreft mogelijke samenhang tussen indicatoren op meerdere aggregatieniveaus en ontwikkelingen in de tijd.

In de afsluitende discussie van dit hoofdstuk zal worden stilgestaan bij de consistentie van resultaten uit de effectiviteitsstudies uit par. 5.4.2 en de resultaten van PISA, zoals besproken worden in deze subparagraaf.

5.4.2.1 Samenvatting van de effecten van systeemvariabelen in PISA (2009, 2012, 2015 en 2016)

De PISA-resultaten worden gerapporteerd voor de kernvariabelen die ook gebruikt zijn om de resultaten van effectiviteitsonderzoek samen te vatten in paragraaf 5.4.2. Hierbij ga ik achtereenvolgens in op de volgende rubrieken: Privéscholen en keuzevrijheid, schoolautonomie en functionele decentralisatie, evaluatie en accountability en stratificatie van scholenstelsels. Vervolgens worden als aanvullende rubrieken behandeld: Financiële randvoorwaarden, het scholingsniveau van leerkrachten, het effect van deelname aan professionele ontwikkelingsactiviteiten, technologische voorzieningen en “curriculum alignment”.

Privéonderwijs en keuzevrijheid

De PISA-resultaten wijzen over het algemeen uit dat verschillen in scores tussen privéscholen en openbare scholen verdwijnen (niet meer significant zijn) nadat gecontroleerd is voor de samenstelling van de schoolbevolking qua sociaaleconomische status (PISA 2012, 2015 en 2018). Competitie tussen scholen heeft geen effect of een effect dat verdwijnt na controle voor sociaaleconomische status (PISA 2009, 2012, 2018).

Schoolautonomie en functionele decentralisatie

In PISA is enkele keren aangetoond dat autonomie in het curriculaire domein een gunstig effect heeft (PISA 2009 en PISA 2012). Hierbij moet worden opgemerkt dat de invulling van “curriculair” hier bepaald is als de vrijheid van docenten om leermethoden en assessment vormen in de klas te bepalen en dus niet in de betekenis van het bestaan van een voorschrijvend nationaal curriculum dan wel een meer “open” curriculum. Van autonomie op het gebied van het zelf managen van financiële en materiële middelen werd vastgesteld dat die een positief effect heeft, maar alleen als dit samengaat met evaluatieve en accountability voorzieningen (gegevens uit PISA 2009 en 2012). In PISA 2015 werd, op basis van een meer algemene autonomie-index, vastgesteld dat deze variabele geen effect had op prestaties.

Evaluatie en accountability

In PISA 2009 werd vastgesteld dat landen die een op standaarden gericht centraal examen hadden gemiddeld 16 punten hoger scoorden bij lezen. De uitkomsten van PISA 2012 lieten een negatief effect zien van het ter beschikking stellen van evaluatiegegevens door scholen aan een “externe autoriteit”. PISA 2015 rapporteert dat het openbaar maken en publiceren van evaluatiegegevens van scholen een

positief effect heeft (gemiddeld een score van 4 punten hoger). In de rapportage over PISA 2018 blijkt dit effect 5 punten te zijn. In Pisa 2018 werd verder een positief effect van assessment en feedback vastgesteld op gelijkheid (equity).

Stratificatie

Voor stratificatie worden twee indicatoren gebruikt: het aantal schoolcategorieën in het voortgezet onderwijs, en de leeftijd waarop voor een bepaalde vorm van voortgezet onderwijs moet worden gekozen. De algemene trend is dat de prestaties lager zijn naarmate er meer schoolcategorieën zijn en er op jongere leeftijd gekozen moet worden. Nederland scoort op beide indicatoren relatief hoog. De positieve effecten van meer geïntegreerde scholenstelsels en latere selectie gaan zowel op wat betreft hogere prestatieniveaus als voor ongelijkheid, maar het meest consistent voor ongelijkheid (resultaten uit PISA 2009, PISA, 2012, PISA 2015 en PISA 2018)

Budget en salarissen

Alleen voor landen met een investering van minder dan 50.000 USD maakt de hoogte van het onderwijsbudget uit voor de prestaties. Bij rijkere landen, cq de OECD-landen is er geen effect (PISA 2018). Uit PISA 2009 gegevens bleek dat er tussen landen een positief effect was voor financiële middelen, salaris inbegrepen, maar dat dit effect binnen landen verdween na controle voor sociaaleconomische status. Salaris bleek in PISA 2018 geen effect te hebben, nadat extreme uitkomsten van 2 landen verwijderd waren.

Het scholingsniveau van leerkrachten

Uit de resultaten van PISA 2012 bleek dat het percentage docenten met een universitaire opleiding geen significant effect op de leerprestaties had. Het percentage bevoegde leraren bleek in PISA 2018 een positief effect te hebben op de prestaties. Verder bleek dat het percentage *full-time* leerkrachten in een land een positief effect te hebben op gelijkheid (equity).

Deelname aan professionele ontwikkelingsactiviteiten

Uit de resultaten van PISA 2012 en 2018 bleek dat van deze deelname geen effect op de prestaties kon worden vastgesteld.

Lerarentekorten

Uit de resultaten van PISA 2012 en PISA 2018 bleek dat door scholen ervaren tekorten een negatief effect op de prestaties hadden.

Curriculum alignment

Curriculum alignment is als zodanig niet systematisch onderzocht in PISA (Ik kwam de term voor het eerste tegen in het rapport over PISA 2018 (Vol V)). De uitkomsten uit PISA 2012 over OTL (gelegenheid om te leren) waren weliswaar spectaculair, maar zijn eenmalig en discutabel. In het rapport van PISA 2018 is er sprake van “a standardized policy for reading related subjects taught at school”; maar de effectiviteit ervan is niet onderzocht.

Directe instructie

In Pisa 2015 werd vastgesteld dat “teacher directed science education” in vergelijking met “enquiry-based teaching practices” een duidelijk positief effect had (het effect van ontdekkend leren was zelfs negatief; OECD, 2016, p 65-73),

5.4.2.2 *Secundaire analyses op PISA 2009 data*

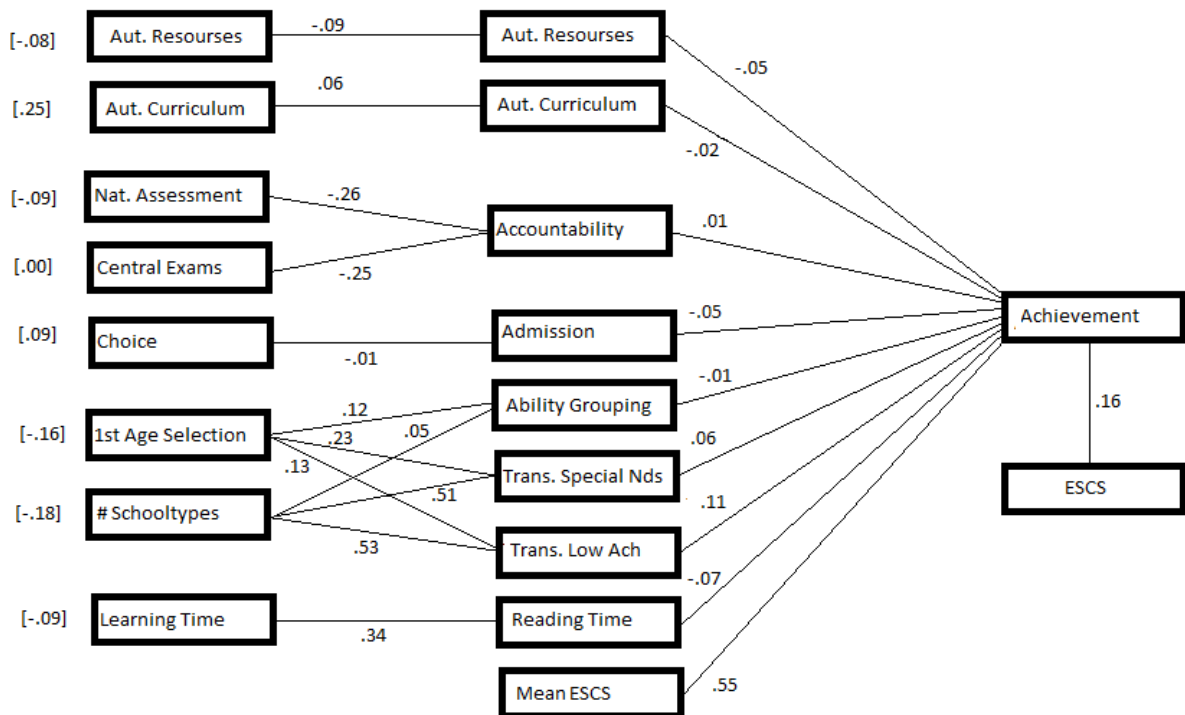
Directe en indirecte effecten van manipuleerbare variabelen op systeemniveau werden onderzocht door Scheerens, Glas et al (2015) in een secundaire analyse van PISA 2009-data. De keuze van de variabelen op systeemniveau werd geleid door de gedachte dat een aantal variabelen zowel een betekenis had op systeem- als op schoolniveau. Daarbij werd uitgegaan van de veronderstelling dat de variabelen op schoolniveau opgevat konden worden als implementatie van de systeemniveauvariabelen op schoolniveau. Het onderzoek werd gezien als een poging om sommige van de assumpties over indirecte causaliteit uit het overkoepelende systeemmodel, zoals gepresenteerd in Hfst. 2, te toetsen.

De data werden geanalyseerd door middel van multi-level covariantie structuuranalyse (Lee, 2007). Het pad-model dat werd geanalyseerd, is weergegeven in de onderstaande figuur.

Een van de opmerkelijkste uitkomsten was de sterke invloed van sociaaleconomische status (SES), en vooral ook de gemiddelde SES op de samenhang van de systeemvariabelen op de leerprestaties. Zonder SES-correctie werd een significant positieve associatie met leerprestaties gevonden van .35. Maar wanneer die correctie wel werd toegepast verdween het verband helemaal. Een overeenkomstige conclusie moest worden getrokken voor het verband tussen autonomie en leerprestaties.

Ongecorrigeerd was de directe associatie tussen schoolautonomie op het vlak van materiele voorzieningen .24 en bij toepassing van de SES- correctie verdween ook hier het verband nagenoeg helemaal. De auteurs wijzen erop dat de correctie voor gemiddelde SES niet onomstreden is en er mogelijk sprake is van over-correctie.

Een tweede opvallende uitkomst is de geringe grootte van de indirecte variabelen (gedefinieerd als het product van de coëfficiënten voor de directe en indirecte associaties). Deze uitkomst falsifieert de hypothese van indirecte causaliteit bij systeemfactoren met een interpretatie op schoolniveau.



Figuur 5.3 geciteerd uit Scheerens, Glas et al (2015). In de figuur worden alle gevonden directe en indirecte relaties van de systeem variabelen met leerprestaties aangegeven, nadat gecorrigeerd was voor sociaaleconomische status van de leerlingen (ESCS). De directe effecten zijn weergegeven in de uiterste linker kolom van de figuur.

Luyten en Scheerens (2014) analyseerden veranderingen in systeemvariabelen in hun relatie tot veranderingen in de gemiddelde leesprestaties voor een subset van landen die deelnamen aan PISA 2009. De voornaamste resultaten waren als volgt:

Er bestond weinig samenhang tussen veranderingen in de systeemvariabelen en veranderingen in de gemiddelde PISA prestaties

Er werd een grote mate van stabiliteit (cq geringe verandering door de jaren heen) geconstateerd voor zowel de systeemvariabelen, als de gemiddelde leestprestaties. Zo was decorrelatie tussen de gemiddelde prestaties over 4 jaargangen van PISA niet minder dan .953.

In de interpretatie van deze uitkomsten wijzen de auteurs een op een groot contrast met opvattingen in rapporten van de OECD (2010) en McKinsey (2010), waarin een rooskleurig beeld wordt geschetst van beïnvloedbare vooruitgang in de prestaties van onderwijsstelsels: *“Our results are generally less encouraging for a reformist educational agenda. Apart from evaluation and assessment practices there was hardly any consistency to be found in the variables that characterized countries that had improved relatively strongly between 2000 and 2009. But, perhaps even more fundamentally, our results underline stability (over a period of 9 years) rather than change in both reading literacy outcomes and associated school and system characteristics. Furthermore, across the board, change scores in system and school characteristics showed small effect sizes, in terms of correlations with changes in reading literacy performance.”*

In een recentere secundaire analyse van PISA-gegevens, die zich uitstrekt over een periode van 15 jaar en 6 afname-tijdstippen, komen Aloisi en Tymms tot overeenkomstige conclusies (Aloisi & Tymms, 2017). Ook zij vonden een opmerkelijke stabiliteit in de gemiddelde prestaties van landen en correlaties, die voor een vergelijking tussen twee jaargangen zelfs opliepen tot .99. Voor zover er veranderingen optraden - de auteurs signaleerden een stijgende trend- konden die grotendeels verklaard worden door veranderingen die niet konden worden gezien als beïnvloedbaar door het onderwijsbeleid, zoals sociaaleconomische en demografische indicatoren. Onderwijsvernieuwing in de zin van curriculumverandering had slechts een zeer klein effect, namelijk .002.

5.4.3 Impressies van enkele facetten van “lopend beleid” en hypotheses over de effecten ervan

Bij de beschrijving van effectiviteit bevorderende systeemkenmerken en processen in par. 5.3.5 werden de volgende facetten van “lopend beleid” genoemd:

- De onderwijsbegroting.
- Nationale beleidsprogramma's die gericht zijn op kwaliteitsverbetering.
- Initiatieven tot curriculumvernieuwing.
- Beleid gericht op het tegengaan van achterstanden en ongelijkheid in het onderwijs.
- De opleiding van leerkrachten.

Zoals werd aangegeven kan op deze facetten van beleid in het kader van deze studie slechts globale aandacht worden besteed. Als het gaat om empirische evidentie over de effectiviteit van deze beleidsfacetten is informatie over de effecten ervan lang niet altijd beschikbaar. Hieronder wordt van elk van de bovenstaande facetten een impressie gegeven van de inhoud, de overeenkomsten met theoretische concepties over werkzame factoren, zoals besproken in eerdere onderdelen van dit hoofdstuk en een enkele keer wordt ook verwezen naar evaluatiestudies.

5.4.3.1 Het onderwijsbudget

Internationaal vergelijkend onderzoek laat zien dat in landen waar de cumulatieve uitgaven per leerling beneden de 50.000 USD liggen extra uitgaven sterk gecorreleerd zijn met leerprestaties. Echter in landen met hoge inkomsten, dus alle OECD-landen, kunnen er grote verschillen in de hoogte van het onderwijsbudget samengaan met gelijke prestatieniveaus. Eveneens kunnen er landen zijn met vergelijkbaar onderwijsbudget, maar sterk uiteenlopende prestaties (OECD, 2016, 42). In welvarende landen lijkt extra investering dus minder het verschil te maken en is meer nodig dan geld om tot excellente prestaties te komen.

Uit Nederlandse studies van het Instituut voor Publieke Sector Efficiëntie Studies (IPSE) uit 2015 en 2019 blijkt dat de productiviteit⁹ van het Nederlandse onderwijs sinds 1980 achteruit is gegaan. (IPSE, 2015, 2019), met name in het p.o., v.o. en mbo. De productiviteitsdaling komt erop neer dat ondanks zeer aanzienlijke verhoging van het onderwijsbudget in de periode van 1980 tot 2018, de prestaties eerder een lichte daling laten zien. Aan de productiviteitsdaling is bijgedragen door de grote vernieuwingsprojecten, geïnitieerd in de jaren 80 van de vorige eeuw, de schaalvergroting en de decentralisatie op zich, met name ook de decentralisatie van de gebouwen. Aan de productiviteitsdaling wordt verder voor een belangrijk deel bijgedragen door budgetverhoging, waarvoor geen productiviteit wordt geconstateerd (de zogenoemde BBP- component). Andere tendensen zijn dat er sprake is van een

⁹ Productiviteit' is een economische term en verwijst naar de verhouding tussen de productie (in het onderwijs: het aantal leerlingen en de toegevoegde waarde per leerling) en de ingezette middelen (kosten) voor die productie. Productiviteit is dus in feite een maat voor de hoeveelheid 'waar' die per euro (vooral belastinggeld) wordt geleverd (IPSE, 2019)

grote productie aan beleid en steeds nieuwe impulsen en dat opbrengst- en evaluatiegegevens schaars zijn. Ook over de effectiviteit van het achterstandsbeleid wordt getwijfeld; deze is niet aangetoond.

5.4.3.2 Nationale beleidsprogramma's gericht op kwaliteitsverbetering

In 2007 formuleerde het Ministerie van OC&W kwaliteitsagenda's voor het primair, algemeen voortgezet en middelbaar beroepsonderwijs (OC&W, 2007, 2007a). De Sectorraden namen medeverantwoordelijkheid voor deze programma's. In elk van deze programma's lag een accent op prestatieverbetering in de basisvakken, taal, rekenen en wiskunde, maar daarnaast werden nieuwe onderwijsopbrengsten beoogd, zoals burgerschapskunde in het voortgezet onderwijs en arbeidsmarktgerichte competenties in het mbo. Alles bij elkaar was er sprake van een breed scala van beoogde eindtermen. In de kwaliteitsagenda's werd melding gemaakt van verschillende hefboomen om tot kwaliteitsverbetering te komen, toetsen, examens en maatregelen om kwaliteitszorg te stimuleren, opleiding en professionele ontwikkeling van leraren, op empirische evidentie gerichte innovatie en een betere aansluiting tussen onderwijsniveaus en schooltypes. De centrale overheid formuleerde de doelstellingen en uitvoeringsvoorwaarden (waaronder de financiering en richtlijnen van "good governance") van deze programma's, terwijl de verdere vormgeving, implementatie en uitvoering gebeurde onder regie van de autonome scholen. In een monitoringstudie van het IVA (Smeets, Vrieze, Driessen en Claassen (2010), werden de volgende indicatoren van de uitvoeringsprocessen beschreven: *onderwijsprocessen* (bv meer tijd ingeroosterd voor lezen en rekenen, formuleren van meetbare leerdoelen, dat echter maar in de helft van de scholen werd opgepakt); *organisatiekenmerken* (bv inzet van externe ondersteuning; verbeterde samenwerking tussen teams); *competenties van de leraar* (verbeteren van klassenmanagement en interactie, analyseren en interpreteren van leerresultaten werd vaak uitbesteed aan interne en externe begeleiders); *opschaling en duurzaamheid* (meer studiedagen en netwerkbijeenkomsten ten behoeve van kennisuitwisseling); *educatieve infrastructuur* (netwerkbijeenkomsten worden van groot belang geacht, meer vraag naar externe ondersteuning); *opbrengstgericht werken* (in 2010 bestond hiervoor volgens de studie van het IVA brede acceptatie; maar er waren problemen met de interpretatie van gegevens en het gebruik hiervan voor onderwijsverbetering en er was nergens sprake van inbedding van opbrengstgericht werken in de organisatie en het management van de scholen). Naast deze "breedtestrategie" in het kader van de kwaliteitsagenda's waren er ook enkele nader gestructureerde deelprogramma's, namelijk "School aan Zet", de "Lerarenagenda" en "Zicht op kwaliteit". Het programma "School aan Zet" probeerde het brede scala aan doelstellingen van de kwaliteitsagenda's vooral te realiseren door middel van korte interventies van vakdidactici en organisatieadviseurs. De externe evaluatie van dit programma uit 2015 wees uit dat de beoogde opbrengsten van het programma, in de zin van verbeterde leeropbrengsten, niet bereikt waren (Van Kuik, e.a., 2015); voor een nader analyse van dit programma, evenals "Zicht op kwaliteit" zij verwezen naar Scheerens, 2016, hfst 12. De onderzoeksprojecten in het kader van "Zicht op kwaliteit" hadden te kampen met problemen om de beoogde experimentele en quasi-experimentele onderzoeksopzetten te realiseren, wat een bekend probleem is in het Nederlandse onderwijs (vgl Scheerens en Doolaard 2013).

De activiteiten die voortkwamen uit de kwaliteitsagenda's bestonden enerzijds uit een op het onderwijs in de klas gerichte kern, waarin op evidentie en toetsing gebaseerde praktijken beoogd werden en anderzijds uit een breed terrein van organisatie en management, professionele ontwikkeling en netwerkvorming (School aan Zet paste in deze laatste categorie). De eerstgenoemde activiteiten (Zicht op kwaliteit, het onderzoeksprogramma "onderwijsbewijs" en het opbrengst gerichte werken) zijn inmiddels afgesloten of naar de achtergrond verdwenen. De organisatieontwikkeling, ondersteuning en netwerkvorming zit duidelijk in de lift. Dat blijkt bijvoorbeeld uit recentere initiatieven van de

sectorraden, zoals het programma Samen Opleiden en Organiseren en aanbevelingen van McKinsey om in de komende 9 jaar miljarden te investeren in de ondersteuning van organisatieontwikkeling in het onderwijs (McKinsey, 2020).

Verdere analyse van de opkomst en ondergang van het opbrengstgericht werken biedt mogelijk verder inzicht in de achtergronden van deze ontwikkeling. Ondanks positieve uitkomsten van effectstudies (Visscher & Ehren, 2011; Visscher, 2020) lijkt de invoering en bredere implementatie van deze werkwijze te stikken.

5.4.3.3 *Initiatieven tot curriculumvernieuwing*

Op de initiatieven in het kader van “Onderwijs 2032”¹⁰ uit 2016 en het latere “Curriculum.nu” zal hier niet nader worden ingegaan, omdat het laatste project in een transitie fase verkeert. Begin 2020 presenteerde de coördinatiegroep van Curriculum.nu de voorstellen van de ontwikkelteams aan de Vaste Kamercommissie OCW.¹¹ Belangrijke thema’s die bij de presentatie ter discussie werden gesteld waren de verhouding tussen vakgebonden onderwijs en meer algemene vaardigheden en de planmatige aanpak van de uitwerking.

5.4.3.4 *Beleid gericht op het tegengaan van achterstanden en ongelijkheid in het onderwijs*

Het onderwijsbeleid gericht op het bestrijden van onderwijsachterstanden heeft een vrij lange traditie. Sinds het begin van de zeventiger jaren van de vorige eeuw werd er eerst onder de noemer van “onderwijsvoorrangsbeleid” en meer recent “onderwijsachterstandenbeleid” extra financiering verstrekt aan scholen met een hoog percentage achterstandsl leerlingen. In de loop der tijd zijn de formules om de steun per school te bepalen enkele keren bijgesteld, maar de basisstrategie veranderde niet. De wijze waarop de extra financiële ondersteuning door de scholen werd ingezet, bleef ook constant door de jaren heen. De middelen werden grotendeels besteed aan klassenverkleining en het inhuren van externe ondersteuning. Reviewstudies onder meer Mulder en Meinen, (2013) en Driessen, (2018) stelden vast dat geen enkel uitgevoerd evaluatieonderzoek heeft kunnen aantonen dat de wijze waarop de scholen de middelen gebruikten effect hadden op verbeterde gelijkheid in leerprestaties. Deze auteurs uitten kritiek op het gebrek aan duidelijke targets voor het achterstandsbeleid en het gebrek aan medewerking van scholen aan het evaluatieonderzoek. De grote autonomie van scholen maakte het evenmin mogelijk om empirisch onderbouwde gestructureerde interventies, bijvoorbeeld naar het voorbeeld van de Amerikaanse “*Comprehensive School Reform Projects*” te implementeren. Ontbrekende positieve resultaten van het achterstandsbeleid zijn dus waarschijnlijk voor een deel niet aangetoond door het ontbreken van optimale onderzoekdesigns. Dat een positieve invloed van het beleid misschien toch aanwezig is geweest, zou kunnen worden opgemaakt uit eerder aangehaalde resultaten in meerdere jaargangen van PISA, waarbij bleek dat de mate waarin de leerprestaties van Nederlandse leerlingen worden beïnvloed door de sociaaleconomische status van leerlingen in het voortgezet onderwijs gematigd is (in de buurt van het OECD gemiddelde); zie ook Scheerens, Timmermans en Van de Werf, 2019).

5.4.3.5 *De opleiding van leraren*

De stelling dat de leraar bepalend is voor de kwaliteit van het onderwijs is een gemeenplaats. In de voorafgaande hoofdstukken werd op verschillende plaatsen gerefereerd aan beschikbare kennis over de vraag hoe groot het leerkrachteffect is. En vooral ook aan leerkrachtkenmerken en instructievormen die dit effect zouden kunnen bepalen. Initiële scholing en nascholing zijn daarom belangrijke hefboomen

¹⁰ file:///C:/Users/Jaap/AppData/Local/Temp/eindadvies-platform-onderwijs2032-ons-onderwijs2032.pdf

¹¹ <https://www.curriculum.nu/voorstellen/>

voor het in stand houden en verbeteren van de kwaliteit van het onderwijs. Hier gaat het dan speciaal om de invloed die de centrale overheid hierbij kan uitoefenen.

In de *Lerarenagenda* (2013- 2020) wordt het verband gelegd tussen de Kwaliteitsagenda's van de overheid en de betekenis van de leraar en de lerarenopleiding. Leren differentiëren en opbrengstgericht werken worden als duidelijke doelstellingen van de lerarenopleiding neergezet. Ook wordt gesproken over een leraar die in staat is "Bildung" tot stand te brengen. Hieronder wordt verstaan "een vorm van reflectie die nodig is om leerlingen culturele bagage mee te geven en hen beter te laten denken" (OCW, 2013, 11). In de "Lerarenagenda" wordt duidelijk gemaakt dat de overheid de autonomie van de sector onderschrijft: "Wij beseffen dat wij deze ambitie niet vanuit 'Den Haag' gaan realiseren. Autonomie van scholen kenmerkt het Nederlandse onderwijsstelsel. Scholen en lerarenopleidingen zijn in eerste instantie zelf verantwoordelijk voor hun kwaliteitsverbetering en daarmee die van hun leraren."

De lerarenagenda beschrijft verder de oprichting van de *Academische PABO'S*, die sinds 2008 moeten zorgen voor een grotere instroom van vwo'ers.

Continue professionele ontwikkeling die moet bijdragen aan een gerichtheid op continue verbetering van het onderwijs worden verder als algemene doelen aangegeven. Een van de instrumenten om hierbij behulpzaam te zijn, is een lerarenregister. Nadruk wordt ook gelegd op organisatieontwikkeling: "Leraar, lerarenteam, leidinggevende en bestuur bouwen samen aan een lerende cultuur waarin continu wordt gewerkt aan verbetering van de onderwijskwaliteit. De leraar heeft daarbij voldoende professionele ruimte, interessante ontwikkelmogelijkheden en carrièreperspectief, zoals specialisatie, onderzoek of coaching. De schoolorganisatie bindt op die manier mensen van binnen en buiten aan het onderwijs. De school werkt als lerende organisatie samen met hogescholen en universiteiten (haar "academische omgeving") en omarmt nieuwe mogelijkheden om het onderwijs te verbeteren, waaronder met betere inzet van ICT" (ibid, p 21) Accreditering van de lerarenopleidingen wordt uitgevoerd door de NVAO. Verder is er sprake van dat de professionele beroepsgroep zelf het nascholingsaanbod valideert en wordt in dit kader samenwerking met de Sectorraden en de Onderwijs Coöperatie genoemd.

In een recenter beleidsdocument van de Vereniging van Hoge Scholen wordt een "Strategische beleidsagenda VH educatieve sector 2019-2024" gepresenteerd (VH, 2019)¹² onder het motto: "Samen toekomstbestendige leraren opleiden". In deze nota wordt de breedte van de leraarsprofessie benadrukt: "De leraar vervult niet alleen een rol als docent, inhoudelijk expert of pedagoog/coach, maar heeft ook een rol op het gebied van bijvoorbeeld innovatie en onderzoek en als begeleider van potentiële nieuwe leraren of startende leraren" (VH, 2019, p 7). In de nota wordt functiedifferentiatie bepleit met betrekking tot leerlinggroepen (leeftijd, speciaal onderwijs) en inhoudelijke specificatie (op een vak gericht, gericht op vakkenintegratie of gericht op pedagogische coaching). Ook worden allerlei vormen van institutionele samenwerking en partnerschappen belangrijk gevonden. De beroepsprofielen moeten op de schop en flexibilisering en maatwerk worden benadrukt. Uitgaande van een (niet nader gespecificeerde) steeds meer divers geachte vraag van het werkveld richt men zich naar bredere kerndoelen, bredere leergebieden en bredere beroepsprofielen. Hierbij preludeert men op ruimte binnen de wet- en regelgeving voor bijvoorbeeld modulair opleiden en het koppelen van civiel effect aan afzonderlijke modules (ibid p 19). Men benadrukt ook de relatie tussen onderwijs en onderzoek en het belang van praktijkgericht onderzoek met het oog op innovatie. Dit aspect van de taakuitbreiding van leraren opent weer een nieuw veld van samenwerkingspartners: "Door het werken met lectoraten heeft het praktijkgericht onderzoek in de educatieve sector een enorme impuls gekregen. Zo zijn er

¹²

https://www.vereniginghogescholen.nl/system/knowledge_base/attachments/files/000/001/088/original/085_040_SBA_EDUCATIE_BW_DEFDEF.pdf?1575577347

werkplaatsen onderwijsonderzoek, (academische) opleidingsscholen, huisacademies en field labs waarin leraren, onderzoekers en andere betrokken professionals met elkaar samenwerken. In januari 2019 is Regieorgaan SIA de call SPRONG educatief gestart die samenwerkingsverbanden helpt zich gezamenlijk te organiseren en te focussen op een thema etc. "(ibid. 30). Op het vlak van de kwaliteitszorg wordt gesproken van een model van "horizontale kwaliteitsborging", waarmee men onder meer het "eigenaarschap" van de lerarenopleiders wil vergroten (p. 39).

Deze visie op de lerarenopleiding roept vragen op over de gerichtheid op het lesgeven van leraren. Het lijkt erop dat de nadruk ligt op allerlei secundaire functies en, zo men wil, gedelegeerde leiderschapstaken, zoals zorgen voor innovatie, externe contacten onderhouden, het doen van onderzoek en collegiale samenwerking. De vraag is niet alleen naar de omvang maar ook naar de diversiteit van aandacht voor deze secundaire functies.

5.4.3.6 Samenvattende impressie van de facetten van lopend beleid

Bij de nationale programma's gericht op de verbetering van de kwaliteit van het onderwijs zijn *evidence based* benaderingen gefocust op het verbeteren van leerprestaties (zoals het opbrengstgericht werken) naar de achtergrond verdwenen.

De decennialange financiële ondersteuning van het onderwijsachterstandsbeleid is vooral gebruikt voor klassenverkleining en externe ondersteuning, waarvan de inhoud en kwaliteit een black box zijn. Extern evaluatieonderzoek heeft geen positieve effecten kunnen aantonen, hoewel uit internationaal vergelijkend assessment onderzoek blijkt dat Nederland, ondanks het sterk gestratificeerde scholenstelsel in het voortgezet onderwijs, "gemiddeld" scoort op door de OECD gebruikte ongelijkheidsindicatoren en het beter doet dan andere landen met een sterk gestratificeerd stelsel. Afgaand op landelijke agenda's bestaat de indruk dat in de lerarenopleiding een steeds groter accent is komen te liggen op secundaire functies, zoals onderlinge en externe samenwerking, programma-ontwikkeling en onderzoek. Incidenteel wordt de vraag gesteld of dit niet ten koste gaat van de primaire onderwijstaak (Van der Ley, 2020)¹³.

Een opmerkelijke conclusie is verder dat aanmerkelijke stijging van de onderwijsuitgaven niet tot prestatieverbetering heeft geleid en dat er sprake is van een verlaging van de productiviteit in het basis- en voortgezet onderwijs (rapporten van IPSE).

5.5 Conclusies

Van de onderzochte structureel ecologische factoren op systeemniveau werden geen of kleine effecten gevonden, zowel in het kader van onderwijs effectiviteitsonderzoek als internationaal vergelijkend assessment onderzoek. Wel zijn er graduele verschillen in effectiviteit tussen de onderzochte factoren. Autonomie en privatisering hadden nauwelijks effect. Wel is er enige onderbouwing voor de effectiviteit van accountability en van de combinatie van autonomie met accountability (new public management). De enige factor met een consistent positief effect is een meer geïntegreerd in tegenstelling tot een sterk gestratificeerd scholenstelsel, met een schoolkeuze op relatief jonge leeftijd. Hoewel het Nederlandse stelsel sterk gestratificeerd is worden de potentieel negatieve effecten gemitigeerd door andere

¹³ "Pak de opleidingen (eindelijk) aan: onderwijs is zorgen dat er geleerd wordt; vaardigheden in lezen vergen (kennis van en vaardigheid in) gerichte instructie, didactiek, leertheorie; zorg dat (toekomstige) onderwijzers weten hoe dat moet!" Presentatie ResearchEd conferentie, 2020
<https://files.sgbg.nl/redeu/uploads/2020/04/21141816/Leij-researchEd-2020.pdf>

kenmerken van de Nederlandse onderwijscontext (zoals aangeven door de OECD, vlg. OECD (2015). Omdat stelselherziening met zeer hoge transactiekosten gepaard gaat, en al eens eerder mislukt is (vgl het Middenschool avontuur in de 70 er jaren), liggen meer incrementele verbeteringen, zoals een verlengde brugperiode en het verbeteren van de mogelijkheden tot doorstroming meer voor de hand.

Als het gaat om potentiële hefboomen van het beleid om iets te doen met deze informatie hangt veel af van het besturend vermogen van de overheid en de bestaande traditie en cultuur in het onderwijs. In theorie is er aanleiding voor een kritisch onderzoek naar mogelijk doorgeschoten autonomie en decentralisatie. De decentralisatie van de werkgeversfunctie naar Sectorraden evenals de dubbelfunctie van deze organisaties wat betreft belangenbehartiging enerzijds en innovatie anderzijds roept vragen op. Wanneer de bestaande autonomie als een niet veranderbaar gegeven wordt gezien is accountability, evaluatie en examinering des te meer van belang. De public managementtheorie en de empirische ondersteuning voor deze vorm van functionele decentralisatie ondersteunen dit standpunt. Instandhouding van het centraal examen en de instandhouding van de eindtoets basisonderwijs zijn daarom van groot belang voor de kwaliteit van het Nederlandse funderend onderwijs.

Tabel 5.8 Hefbomen voor beleid gericht op systeemeffectiviteit

Besturingsmechanisme	Gemeten dimensies en effecten	Hefbomen
Synoptische planning	-	Specifiekere doelstellingen en eindtermen ook voor de lerarenopleiding
Marktwerking/autonomie	Geen duidelijk effecten van autonomie en privéonderwijs	Nauwelijks "aanraakbaar", monitor kosten en baten
Evaluatie gestuurde planning	Wisselende en geringe effecten van accountability en evaluatie	Instandhouding centrale examens en eindtoets
New public management	Enige ondersteuning voor combinatie procesautonomie outputcontrole	Idem, instandhouding inspectie en accountability. Retroactieve invulling van curriculum-alignment
Stelselherziening (adaptieve planning)	Brede empirische onderbouwing voor minder stratificatie	Incrementele stappen zoals verbrede brugperiode

In de evaluatie van het Nederlandse onderwijs door de OECD (Nusche, 2016) werd de aanbeveling gedaan om te komen tot het formuleren van duidelijker onderwijsdoelstellingen. Heldere doelstellingen en standaarden zijn verder van groot belang voor zowel de ontwikkeling van curricula als examens. Men zou zelfs, na een onderbouwde keuze en uitwerking van doelstellingen, kunnen volstaan met een eventueel bijgesteld examen en het verder kunnen overlaten aan de scholen, uitgevers en ander marktpartijen om op het examen gerichte leerlijnen en volgsystemen te ontwikkelen. De beschrijving van de ontwikkelingen rondom de lerarenagenda en beleidsplannen van deze sector suggereert dat de aanbeveling van de OECD zeker ook betrokken moet worden op de lerarenopleidingen.

Ontwikkelingen rondom de lerarenopleiding en de op kwaliteit gerichte programma's van de overheid, besproken als voorbeelden van "lopend beleid", wekken de indruk dat er sprake is van een afname van de belangstelling voor op evidentie gebaseerde werkwijzen, zoals "opbrengstgericht werken" en weerstand tegen toetsing. Daar tegenover staan een toenemende focus op secundaire functies in het onderwijs zoals "governance" leiderschap en samenwerking en uitbreiding van het beroepsprofiel van

de leraar met taken op het terrein van onderzoek, ontwikkeling en deelname aan overlegcircuits¹⁴. Hier doet zich vooral de vraag voor naar de proportionaliteit; zoals eerder aangegeven is er empirische steun voor de effectiviteit van taakgerichte samenwerking tussen docenten. Ook is investering in nascholing bij het gebruiken en interpreteren van toetsen zeer goed te verdedigen, maar het roept vragen af of leerkrachten minder les zouden moeten geven om aan onderzoek te doen.

Bij de curriculumvernieuwing blijft de aandacht voor kwalificatie, in verhouding tot sociaal emotionele facetten en “persoonsvorming” een thema dat van groot belang is als het gaat om de prioriteiten in het onderwijsbeleid wat betreft het borgen en verbeteren van basiskwaliteit.

¹⁴ In een recente missive (maart 2021) pleit de Onderwijsraad bijvoorbeeld voor meer structureel geld “zodat leraren gemiddeld minder uren voor de klas hoeven te staan. De tijd die vrijkomt, kunnen ze dan besteden aan het ontwikkelen van onderwijs. Het is zaak dat leraren lessen, lesmateriaal en toetsen ontwerpen en implementeren, en ook analyseren en evalueren.” <https://nos.nl/artikel/2374025-onderwijsraad-basisschoolleraar-moet-zich-meer-kunnen-richten-op-onderwijs.html>

6 HEFBOMEN ALS SCENARIO'S VOOR VERBETERING EN HET TEGENGAAN VAN INEFFECTIVITEIT

6.1 Inleiding

In dit hoofdstuk wordt om te beginnen een schematische samenvatting gegeven van de uitkomsten van de inventarisatie van wat er werkt op micro-, meso- en macroniveau en van de theoretische interpretatie van die uitkomsten. Daarbij wordt ook ingegaan op condities en mechanismen die minder effectief of zelfs contraproductief zijn. De resultaten over effectiviteit en ineffectiviteit worden vergeleken met een tweedeling, die in de internationale literatuur wordt gebruikt onder de noemer van “progressief/traditioneel” (Peal, 201) of “progressief/formalistisch” (Guthrie, 2011).

Omdat in dit rapport de besturing van het onderwijs door de overheid de centrale invalshoek is, volgt een beschouwing over het besturend vermogen van de overheid in de Nederlandse context. Daarin wordt gebruik gemaakt van een onderverdeling uit de besturingstheorie in doelsturing, adaptieve sturing en routine sturing. Doelsturing opent de vraag naar de normatieve uitgangspunten van het beleid. Adaptieve sturing kijkt naar de mogelijkheden van structurele hervormingen en routine sturing naar de diverse facetten van wat eerder is aangeduid als lopend beleid. Het besturend vermogen van de overheid speelt mee in de afweging over de haalbaarheid van mogelijke beleidsscenario's om het niveau en de gelijkheid van het onderwijs positief te beïnvloeden.

6.2 Samenvatting van empirische resultaten

In de onderstaande tabel (6.1) worden de resultaten voor effectiviteit bevorderende factoren op het microniveau samengevat.

Tabel 6.1 Wat wel en niet werkt op het microniveau.

Typering	Effect grootte	Conceptuele basis	Ineffectieve tegenhanger
“Pedagogical content knowledge”	-	Didactische analyse	Relativering van het belang van kennis
Gelegenheid om te leren (OTL)	.30 *	Een goed uitgelijnd curriculum	De mythe van inhoudsloze vaardigheden
Klassenmanagement van effectieve leertijd	.20	Carroll model	
Directe instructie	.50	Cognitieve informatieverwerkingstheorie	“Open”, ongeleide, op constructivisme geïnspireerde instructie,
Vakgebonden leerstrategieën en meta-cognitie	.40	Doelstellingstaxonomieën Cognitieve informatieverwerkingstheorie	
Ordelijk klimaat	.30	Formele regulering en bureaucratie	Ontbreken van regels en orde “over-permissive”
Evaluatie en feedback	.20	Cybernetisch principe	Anti-meet lobby

* De weergegeven coëfficiënten zijn voor het merendeel gebaseerd op afrondingen van de coëfficiënten uit de eerste kolom van tabel 3.3. De uitkomst over direct teaching is gebaseerd op de meta-analyse van Stockard et al., 2015, en de coëfficiënt voor Evaluatie en feedback is naar boven afgerond op basis van aanzienlijk hogere waarden gevonden in de meta-analyses van Hattie (2009). De effectgrootte voor OTL is bijgesteld op basis van de analyses van Scheerens et al. (2017)

De aangegeven effectgroottes zijn schattingen, waarbij soms de effecten van sub variabelen gemiddeld zijn, zoals bij evaluatie- en feedback. Een meer gedetailleerdere weergave wordt gegeven in Hoofdstuk 3. De categorie “vakgebonden leerstrategieën en meta-cognitie” staat voor gestructureerd onderwijs gericht op hogere cognitieve processen, zoals transfer van het geleerde naar nieuwe gebieden en inzicht (vgl. Drost & Verra, 2019). In de “Toolkit” van de Engelse Educational Endowment Foundation, komt deze typering overeen met de daar gegeven invulling van “self-regulated learning”. Als zodanig dus een vorm van “directe instructie” en anders dan het constructivistische “independent learning”.

In tabel 6.2 wordt een samenvatting gegeven van de resultaten voor het mesoniveau (de schoolorganisatie).

Tabel 6.2 Wat wel en niet werkt op het schoolniveau.

Typering	Effectgrootte	Conceptuele basis	Ineffectieve tegenhanger
Schoolvariabelen die staan voor aggregaties van condities op microniveau (OTL, schoolcurriculum, leertijd, huiswerk en prestatiegerichtheid)	.30 *	Professionele bureaucratie; leerkracht autonomie Rationeel geleide school	Autocratisch leiderschap, laissez faire beleid
Schoolvariabelen die uitdrukking geven aan directe ondersteuning van het primaire proces. Instructiegericht leiderschap, duidelijke schoolregels en taakgerichte samenwerking	.20	Model van onderwijskundig leiderschap en de lerende organisatie, waarin evaluatie en feedback het centrale mechanisme zijn	Accent op secundaire condities en niet essentiële taakaspecten van leraren (zoals onderzoek, en curriculumontwikkeling)
Algemene organisatie-factoren die gericht zijn op het welzijn van leerkrachten en het onderhouden van externe relaties; transactioneel leiderschap	.10	De school als open systeem en als sociale werkgemeenschap.	Gevaar van overaccentuering die ten koste gaat van de aandacht voor het primaire proces.

* De vermelde gegevens zijn gebaseerd op samenvoeging en uitmiddeling van uitkomsten die vermeld zijn in Tabel 4.5. Het onderscheid in beredeneerde effectgroottes tussen direct ondersteunende en meer algemene schoolorganisatiekenmerken is geïnspireerd op enkele afzonderlijke onderzoeken en is in de meta-analyses, samengevat in tabel 4.5, niet gemaakt. De waarden in de tweede en derde rij, tweede kolom zijn daarom hypothetisch.

Deze tabel bouwt op de gegevens over theorieën van de school, zoals samengevat in tabel 4.3, en effectgroottes, vermeld in tabel 4.5.

De logica van deze driedeling is het principe dat factoren dicht bij het primaire proces meer invloed hebben en dat de efficiencywinst van leerkrachtautonomie niet gesmoord moet worden in bestuurlijke drukte en verandering om de verandering.

In het onderstaande overzicht wordt een beeld gegeven van de invloed van systeemfactoren (macroniveau). Omdat op dit niveau nauwelijks meta-analyses zijn uitgevoerd, wordt de schatting van uitkomsten weergegeven als “geen”, “geen tot zeer klein effect”, “klein effect” en een “betekenisvol effect”. De tabel is gebaseerd op gegevens vermeld in Hoofdstuk 5 en de Annex.

Tabel 6.3 Samenvattende Impressie van de invloed van factoren op systeemniveau

Typering	Indruk van effect	Conceptuele basis	Ineffectieve tegenhanger
Vrije schoolkeuze & privatisering	Geen effect	Economische theorie, marktwerking	
Autonomie	Geen tot zeer klein effect	Minder regeldruk, meer responsiviteit	
Accountability	Klein effect	<i>Public choice</i> theorie; prestatieprikkel; theorie over curriculum-alignment	Immunisering tegen externe evaluatie
“New public management”	Klein effect	Functionele decentralisatie	Geen of “zachte” opbrengstmetingen. Anti-opbrengstgerichtheid en verheffing van middelen en processen tot doelen.
Stratificatie	Betekenisvol negatief effect	Te vroege selectie; onvoldoende mogelijkheden tot doorstroming tussen schoolcategorieën	In dit geval is er een effectieve tegenhanger, namelijk meer geïntegreerde schoolstelsels
Verhoging onderwijsbudget	Alleen effectief in arme landen	Effect van betere personele en materiële voorzieningen	Verminderende meeropbrengst en ondoelmatigheid van bestedingen
Opleidingsniveau en professionele ontwikkeling leraren	Geen effecten in OECD- landen	Verwacht positief effect van beter opgeleide leraren	
Lerarentekort	Klein negatief effect	Groot belang van voldoende personeelsbezetting	
Curriculum alingment	Klein effect van OTL	Positieve verwachtingen over verticale en horizontale “uitlijning” van het curriculum	Losse koppeling en autonomie in het curriculum domein

De resultaten in tabel 6.3 berusten in belangrijke mate op analyses gebaseerd op PISA en moeten met de nodige voorzichtigheid geïnterpreteerd worden. Gezien het correlatieve en cross-sectionele karakter van het onderzoek wekt het spreken van “effecten” een suggestie van causaliteit die strikt genomen niet gepermitteerd is. Dit laatste geldt overigens ook voor de resultaten over schooleffectiviteit (mesoniveau). Alleen op het microniveau is er een deel van het onderzoek dat gebaseerd is op experimenteel en quasi-experimenteel onderzoek. Wanneer we de invloed van factoren op micro- meso- en macroniveau vergelijken is een bekend patroon dat de effecten op microniveau groter zijn dan de effecten op meso- en macroniveau en dat de effecten op mesoniveau groter zijn dan de effecten op macroniveau. Dat blijkt ook in het hier gepresenteerde overzicht.

6.3 Verklarende mechanismen voor effectiviteit en ineffectiviteit

De achterliggende concepties en theorieën, die inzicht geven in de vraag waarom bepaalde factoren beter werken dan andere, zijn eveneens een onderdeel van de kennisbasis over onderwijseffectiviteit. “Directe instructie” en gestructureerd onderwijs wordt ondersteund door de cognitieve informatieverwerkingstheorie. Relevante optimaliserende condities zijn: de erkenning van het belang van vakgerichte kennis, didactisch goed geschoolde leerkrachten, formatieve en summatieve toetsing en feedback.

De belangrijkste taak van de school is om deze optimaliserende condities zo goed mogelijk te waarborgen. Dit komt tot uitdrukking in voldoende leertijd en “opportunity” om vakgerichte kennis te verwerven en een stijl van leiding geven die de autonomie van leerkrachten respecteert, maar ook taakgericht ondersteunt.

De belangrijkste taak van de overheid is om de kwaliteit van het onderwijs te stimuleren en te waarborgen, met name door duidelijke onderwijsdoelstellingen te formuleren, zorg te dragen voor een goed uitgelijnd systeem van examens en toetsen en te zorgen voor inspectie waarbij scholen verantwoording afleggen over hun prestaties. *Last but not least* ook het zorgen voor een lerarenopleiding van hoog niveau, waarin leraren vooral worden opgeleid om les te geven en een zekere regulering op de marktwerking bij dienstverlening en ondersteuning.

De voornaamste bron van *ineffectiviteit* op het microniveau is onvoldoende gestructureerd “open” onderwijs, waarbij te snel een beroep wordt gedaan op zelfregulatie. Op het mesoniveau zijn dat gebrek aan taakgerichte focus, overaccentuering van secundaire organisatie-condities en niet-essentiële functies van leraren, zoals het doen van onderzoek en curriculumontwikkeling. Problemen op het macroniveau zijn erosie van checks & balances, bij het functioneren van het georganiseerde onderwijsveld, bijvoorbeeld dooraantasting van de eindtoets en het centraal examen, en overproductie van diensten en netwerken gericht op innovatie, scholing en begeleiding.

In de volgende paragraaf wordt een nadere interpretatie van deze uitkomsten gegeven, door ze in verband te brengen met de tweedeling van traditioneel en progressief onderwijs.

6.4 “Progressively worse” (Peal, 2014)

In dit rapport is niet alleen geprobeerd om aan te geven wat effectieve methoden en strategieën zijn, maar ook wat niet of minder effectief is. Door andere auteurs wordt een tweedeling gemaakt die meer politiek en normatief klink, namelijk de tweedeling tussen progressief en traditioneel (Peal, 2014) of tussen progressief en formalistisch (Guthry, 2021). Deze manier van labelen heeft zijn nadelen, omdat ze sterk polariserend werkt en minder oog heeft voor tussenposities en misschien verwarring scheidt. Voorbeelden van nuancerings bij de tweedeling zijn het beschrijven van een geleidelijke overgang van sterk geleide instructie naar lichtere vormen ervan (*fading*) en het gebruik van directe instructie voor het aanleren van vakgebonden leerstrategieën en meta-cognitie (zie Hoofdstuk 3). De aanduiding traditioneel/progressief kan verwarring opwekken, omdat hij niet zonder meer samenvalt met conservatief en progressief in politieke zin.

Maar ook bij deze auteurs staat de vraag naar effectieve en minder effectieve praktijken centraal. Voorafgaand aan een vergelijking met de hier getrokken conclusies een smaakbepalend citaat van Robert Peal:

“Moeten kinderen leren van de wijsheid van een gezaghebbende leraar, of moeten ze onafhankelijk leren en dingen zelf ontdekken? Moeten leerlingen een academisch, vakgericht curriculum leren of moet dat worden gezien als het alleen maar het vullen van hun hoofd met “slechts kennis”, waar “vaardigheden” bruikbaar zijn? Moeten leerlingen gestimuleerd worden door de structuur van beloning en examens of worden zij geacht gemotiveerd te worden door lessen die “relevant” en “leuk” zijn. Moeten kinderen op hun vingers getikt worden voor wangedrag en niet werken of is dat wreed en autoritair?”. Het is duidelijk dat dit voor Peal retorische vragen zijn. Guthry, (2021) baseert zijn tweedeling op een uitgebreide documentatie over de effectiviteit van onderwijs in ontwikkelingslanden. Zijn typering van “progressief” onderwijs komt sterk overeen met die van Peal, maar de tegenhanger ervan, die hij “formalistisch” noemt, sluit dicht aan bij traditionele onderwijspraktijken in ontwikkelingslanden, die sterk bepaald worden door frontaal en klassikaal lesgeven. In de onderstaande tabel wordt een overzicht gegeven van de kenmerken van progressief versus traditioneel en formalistisch in het werk van Peal en Guthry.

Tabel 6.4 Overzicht van de typering “progressief” versus “traditioneel” en “formalistisch” onderwijs

Peal		Guthry	
Progressief	Traditioneel	Progressief	Formalistisch
Kindgericht , “open” onderwijs	Directe instructie	Kindgericht “open” onderwijs	Leerkracht geleid onderwijs
Kennis staat niet centraal, maar vaardigheden	Vakgericht, kennis gaat vooraf aan algemenere vaardigheden	Competentiegericht onderwijs	Een expliciet vakgericht curriculum
Strikte discipline en morele opvoeding zijn uit den boze	Duidelijke gedragsregels en een ordelijk klimaat	Liberaal en democratisch klimaat	Een duidelijke hiërarchie in de schoolorganisatie
Het succes van leerlingen wordt door het sociaaleconomische milieu bepaald	Deze opvatting wordt gezien als een excuus om niet het maximale uit alle leerlingen te halen. Hoge verwachtingen en uniforme standaarden		
		In ontwikkelingslanden bestaat de neiging om traditioneel en formalistisch te vervangen door progressief (vaak om steun van externe sponsors te verwerven)	Traditioneel onderwijs wordt extern ondersteund door een bijpassende lerarenopleiding, syllabi, inspecties en examens

Het werk van beide auteurs documenteert de effectiviteit van directe instructie en “traditioneel” onderwijs en bevat convergerende typering van ineffectief “progressief” onderwijs. Guthry beschrijft het falen van het progressieve model in ontwikkelingslanden in een uitvoerig oeuvre (Guthry, 2011, 2021). Peal beschrijft de geschiedenis van de onderwijsvernieuwing in Engeland sinds de zeventiger

jaren van de vorige eeuw. Hij herleidt de wortels van het progressieve onderwijs naar Rousseau, Dewey, Piaget en Vygotsky en de Amerikaanse positieve psychologie. Zijn historische beschrijving begint bij een tijdelijke opleving van gestructureerd onderwijs in de vorm van de Numeracy en Literacy programma's in de tweede helft van de negentiger jaren. Deze programma's waren succesvol, maar in de terminologie van Peal nam in de erop volgende jaren onder het bewind van New Labour de "progressieve orthodoxie" de macht weer snel over. Opmerkelijk ook was de explosie van het budget voor onderwijsondersteunende functies op het terrein van innovatie, "kwaliteit" en schoolleiderschap in de periode van New Labour, begin van de 20^{ste} eeuw, die samenviel met een terugval van de prestaties. Vage onderwijsidealen bleken een goed verdienmodel voor bestuurders, schoolleidersorganisaties en onderwijsondersteuners. De positieve noot waarmee Peal eindigt is zijn optimisme over een nieuw opgerichte schoolsoort in Engeland, de zogenoemde "academies", scholen die minder van de overheid afhankelijk zijn. In een aantal van deze academies wordt strak en traditioneel onderwijs gegeven (de Londense school uit de TV serie Klassen) is hiervan een voorbeeld.

Zowel uit het werk van Peal en Guthry als uit het in hoofdstuk 3 geciteerde artikel van Stockard et al. (2018) spreekt verbazing over het feit dat, ondanks de sterk onderbouwde bewijskracht van effectiviteit van directe instructie, al decennialang de voorkeur wordt gegeven aan "progressief" onderwijs.

Reden om iets nader in te gaan op mogelijke verklaringen voor de aantrekkelijkheid van "progressief" onderwijs. Om te beginnen wordt een nadere typering gegeven.

Robert Peal beschrijft in zijn boek het summum van het door hem bekritiseerde progressieve onderwijs. Het ging om een nieuw curriculum ontwerp dat in 2004 werd ingediend door de toenmalige Inspecteur-Generaal, Mike Tomlinson. Dit nieuwe curriculum werd gekenmerkt door het afschaffen van gestandaardiseerde examens en de vervanging daarvan door modulaire (cq maatwerk) diploma's, volledig zelfgestuurde progressie van leerlingen en verplichte lessen in communicatieve vaardigheden en "family responsibilities" (Peal, 2014, p 141). Hij eindigt met de verzuchting dat het rapport gelukkig door de Minister van Onderwijs naar de prullenmand werd verwezen, maar toch bijna was aangenomen.

Een eerste verklaring voor de populariteit van progressief onderwijs is daarom simpelweg dat het kennelijk aantrekkelijk gevonden wordt door zowel aanbieders als "afnemers". Het imago van progressief heeft kennelijk ook een goede pr-waarde, die onder meer te maken heeft met een kindvriendelijk imago, verbinding met ICT en moderniteit. Zie ook de bespreking van "progressively worse" van Paul Kirschner (2020). Personaliserend, of gepersonaliseerd leren, is de meest recente typering van progressief onderwijs. In een brochure van Noordhof (ongedateerd)¹⁵ wordt gepersonaliseerd leren beschreven als een combinatie van vraaggestuurd leren, competentiegericht¹⁶ onderwijs en op maat lesgeven. Gepersonaliseerd leren onderscheidt zich van eerdere typering van adaptief onderwijs, zoals individualisering en differentiatie door de grotere autonomie en "verantwoordelijkheid" die de leerlingen zouden moeten krijgen. De leerling wordt gezien als een zelfsturende lerende, die weet wat noodzakelijke en additionele kennis is. De docent wordt in deze typering niet genoemd, zelfs niet als coach.

Voor deze opvatting over zelfsturend leren bestaat een ideologische basis. *Een tweede antwoord* op de vraag naar het waarom van de populariteit van "progressief onderwijs" heeft daarom te maken met

¹⁵ Gepersonaliseerd leren. Wat is het en waar komt het vandaan? *Persoonlijk.Noordhoff.nl* Aangekend moet worden dat er uiteenlopende interpretaties van "gepersonaliseerd" onderwijs bestaan. Voor een meer "gestructureerde" invulling, gebaseerd op adaptief toetsen, zij verwezen naar Veldkamp, 2019

¹⁶ Gerefereerd wordt aan 21st century skills.

bijdragen uit de geesteswetenschappelijke pedagogiek en de weerklank die dit vindt in het onderwijs. Het ideaal van de leerling die “eigenaarschap heeft over zijn eigen leerproces” is terug te voeren op wat Duijker in 1976 analyseerde als “de ideologie van de zelfontplooiing”. Opmerkelijk is dat Duijker’s artikel een kritische reactie was op de bekende “Contouren nota” van Minister van Kemenade. Duiker richtte zijn pijlen op de daar geventileerde idealen van ontdekkend leren, brede vorming in combinatie met het afschaffen van rapporten (in de klassieke betekenis) en examens. Precies dezelfde kenmerken worden momenteel centraal gesteld in geesteswetenschappelijk pedagogische theorieën, zoals de “Non-affirmative education theory” van Ujens en Ylimaki (2017). In deze theorie worden interventies en sturing in het onderwijs al gauw geassocieerd met onwenselijk autoritarisme. In plaats daarvan worden eisen aan het onderwijsproces gesteld in de geest van opvoedbaarheid, erkenning van de leerling en uitnodiging tot zelf werkzaamheid. Biesta (2018) voegt hieraan het ideaal van “subjectificatie” toe, dat in de interpretatie van de Onderwijsraad geassocieerd is met persoonsvorming. Deze pedagogische inbreng resoneert met afwijzing van rendementsdenken en verzet tegen “toets-gekte”.

Behalve voorkeuren van aanbieders en klanten en een pedagogisch ideaal zijn er ook commerciële belangen in het geding.

Zoals hierboven aangegeven staat de docent bij gepersonaliseerd leren op afstand en wordt zelfs niet genoemd als coach. In de inmiddels ontstane praktijk van gepersonaliseerd leren in de VS wordt de rol van ondersteuning overgenomen door de technologie. In een evaluatierapport van NEPC (National Education Policy Center, 2019) wordt de volgende beoordeling van door technologie ondersteund gepersonaliseerd leren gegeven:

“Our analysis reveals questionable educational assumptions embedded in influential programs, self-interested advocacy by the technology industry, serious threats to student privacy, and a lack of research support” (ibid. 3). Commerciële belangen van onder meer technologiegiganten zijn een *derde antwoord* op de vraag naar de aantrekkelijkheid van deze meest recente manifestatie van progressief onderwijs. De commerciële betekenis van de “soft skills”-beweging speelt hierbij eveneens een rol belangrijke rol. In kritische commentaren hierop van Effrem & Robbins, (2019) en Williamson (2019) wordt gewezen op bedreigingen van de privacy van leerlingen, bij “affective computing” en daaraan verbonden “embedded assessment” (Scheerens, Van der Werf en De Boer, 2020).

Een vierde invalshoek biedt een verklaring van de institutionalisering van “progressief onderwijs” als een overlevingsstrategie voor organisaties in de (semi-) publieke sector. De “public choice”-theorie (Niskanen, 1971) beschrijft de neiging van dit soort organisaties tot een uitbreiding van het dienstenpakket als doel op zich. De ruimte daartoe is des te groter naarmate de missie van deze organisaties, bijvoorbeeld “innovatie”, algemener en vager is en naarmate niet alleen “de tucht van de markt” ontbreekt maar er ook geen rekenschap (accountability) moet worden afgelegd. De hypothese dat uitdijende netwerken, overlegplatforms en samenwerkingsverbanden in het Nederlandse onderwijs kunnen worden toegeschreven aan dit soort deels verborgen organisatie-mechanismen kan verder worden onderbouwd op basis van inzichten uit de organisatiekunde en politicologie (bv Warner & Havens, 1968, Masuch e.a., 1998 en Allison, 1969, 1971).

6.5 Over vernieuwing en verbetering; het besturend vermogen van de overheid

De kennis over onderwijseffectiviteit die in deze studie is gepresenteerd, kan worden gezien als een geheel van inzichten dat gebruikt kan worden om het onderwijs te vernieuwen of te verbeteren. Van vernieuwing is sprake wanneer zowel de doelen als de middelen worden veranderd; bij verbetering

worden de doelen als gegeven beschouwd. Verder is de overheid gekozen als de actor die deze besturing ter hand neemt. Met behulp van de besturingstheorie (De Leeuw, 1990, Kramer, 1978, Kickert, 1979), specifiek het besturingsparadigma, ontwikkeld door De Leeuw, kan ons thema “hoe het onderwijs te verbeteren op basis van de kennis over onderwijs-effectiviteit” meer systematisch als een besturingsprobleem worden behandeld.

Volgens het besturingsparadigma bestaat een besturingssysteem uit een bestuurder, aangeduid als het Besturend Orgaan (BO), een bestuurd systeem (BS) en een omgevingscomponent (E – voor “environment”). Hier gaat het om een besturingssituatie waarin de rijksoverheid wordt gezien als BO en het funderend onderwijs als bestuurd systeem BS. Als omgeving kan de onderwijscontext (vervolgonderwijs) en de ruimere maatschappelijke context worden gekozen. Het besturingsparadigma bevat verder een aantal concepten waarmee de besturing nader wordt gecategoriseerd en gekwalificeerd. Allereerst worden voorwaarden voor effectieve besturing onderscheiden en vervolgens wordt besturing onder meer getypeerd als “routine sturing”, “adaptieve sturing” en “doelsturing”.

De voorwaarden voor effectieve besturing zijn als volgt:

- 1) Het BO moet een doel specificeren dat het BS moet realiseren en dat dient als richtsnoer bij de besturing. In ons geval is dat het verbeteren van het niveau en de gelijkheid van onderwijsprestaties van leerlingen.
- 2) Het BO moet over een model van het BS beschikken. Dat model is hierbij gegeven met het hiërarchische CIPO- model (Fig. 2.3) en de invulling ervan met de effectiviteit bevorderende factoren op micro- meso- en macroniveau in de rest van dit rapport.
- 3) Het BO moet beschikken over informatie die betrekking heeft op de toestand van de in het model opgenomen factoren en de daarop inwerkende omgevingsgrootheden. In ons voorbeeld kan hierbij worden gedacht aan informatie over de beginsituatie van een overwogen sturingsingreep, in de zin van een peiling van de onderwijsprestaties op dat moment, een beeld van de invloed van omgevingsfactoren, zoals bijvoorbeeld de sociaaleconomische status van leerlingen en de vigerende invulling van de beheersbare procesinstellingen, zoals bijvoorbeeld de voorkeur voor “progressief” onderwijs in Nederland.
- 4) Het BO moet beschikken over voldoende besturingsvariëteit.

Bij voorwaarde 4 is de aanname dat het BO überhaupt beschikt over daadwerkelijk te implementeren stuurmaatregelen en verder dat die voldoende variëteit hebben om effectief te kunnen interveniëren. De eerste vraag is fundamenteel en heeft betrekking op het stuurvermogen (of besturend vermogen) van het BO. Het besturend vermogen kan verminderd worden door een viertal condities:

- a) De mate van acceptatie van de doelstelling. Door eigen doelgericht gedrag van het BS, en eventuele alternatieve doelen die vanuit de omgeving worden gesteld, wordt dit bemoeilijkt. In onze situatie zou een voorbeeld kunnen zijn dat actoren in het onderwijsveld, zoals de Sectororganisaties, meer accent willen leggen op sociale en persoonsgerichte vorming dan de overheid.
- b) Een minder adequaat model van het BS, waarover het BO beschikt. Dit is zeker van toepassing op onze besturingssituatie. De empirische evidentie voor het hier gepresenteerde effectiviteitsmodel heeft zijn beperkingen (zie het Voorwoord van dit rapport) en is op sommige onderdelen zwakker dan op andere onderdelen (zie ook de eerder in dit hoofdstuk gegeven samenvatting).
- c) De informatieverwerkingscapaciteit van het BO. Ook deze staat ter discussie, niet voor niets wordt gedacht aan verbetering door middel van een Kennisinstituut¹⁷.

¹⁷ Hier is veel meer over te zeggen, onder meer over de groei van het praktijkgerichte onderzoek en recent het door Scheerens en Van der Werf (2020) bekritiseerde McKinsey rapport.

d) De besturingsvariëteit.

Om deze laatste conditie nader te verklaren worden potentiële stuurmaatregelen gecategoriseerd als routinesturing, adaptieve sturing en doelsturing¹⁸.

Routinesturing heeft betrekking op wat in hoofdstuk 5 is behandeld onder de noemer van “lopend beleid”, het zijn maatregelen waarbij geen bijstellingen van doelen of structurele hervorming in het geding is.

Adaptieve sturing heeft betrekking op structurele bijstelling van het BS; een voorbeeld zou zijn het opnieuw oprichten van een “Middenschool” voor de eerste fase van het voortgezet onderwijs.

Bij doelsturing worden de doelen bijgesteld, mogelijk onder invloed van veranderde maatschappelijke omstandigheden.

Twee onderdelen van de zojuist besproken besturingstheorie worden gezien als behulpzaam bij de vraag naar de toepassing van de kennis over onderwijseffectiviteit. Enerzijds de zojuist onderscheiden typen van sturing en anderzijds het besturend vermogen van de overheid.

Het onderscheid tussen routinesturing, adaptieve sturing en doelsturing plaatst twee belangrijke facetten van het functioneren van het onderwijssysteem op de voorgrond. Doelsturing zet de kwestie van mogelijk verschillende inhoudelijke en normatieve visies op onderwijs en op de kwaliteit van het onderwijs op de agenda. Adaptieve sturing vestigt de aandacht op de complexe opbouw van het Nederlandse onderwijsstel, zowel wat betreft het scholenstelsel als het bestuurlijke middenveld.

Discussie over en onderzoek naar doelstellingen

Adviezen van de Onderwijsraad over de conceptualisering van onderwijskwaliteit en discussies over curriculumvernieuwing geven aan dat er momenteel geen duidelijkheid is over de onderlinge prioritering van vakgericht onderwijs enerzijds en sociale en persoonsvorming anderzijds. Gedeeltelijk gaat het hier om een ouder debat over de school als onderwijsinstelling of vormingsinstituut. Ook in het geding hierbij is de internationale aandacht voor “soft skills”, evenals vragen over de onderwijsbaarheid en meetbaarheid hiervan (Vgl. Scheerens, Van der Werf en De Boer, 2020). Keuzes in deze hebben betekenis voor de opbouw van curricula (bijvoorbeeld de verhouding tussen onderwijstijd voor de verschillende vakgebieden en overige activiteiten) en voor de inhoud van examens. Er is alle aanleiding om te stellen dat de overheid “aan zet” is om zo’n nationaal debat, onderbouwd door een grondig empirisch doelstellingsonderzoek, te organiseren.

De macht van het bestuurlijke middenveld; is adaptieve (bij)sturing nodig en uitvoerbaar?

Traditioneel wordt het Nederlandse onderwijssysteem gekenmerkt door invloedrijke organisaties, de vroegere koepelorganisaties en de huidige Sectorraden. (Bronneman-Helmers, 2011). In de huidige Corona crisis komt dit typisch Nederlandse verschijnsel ook in andere sectoren onder vuur te liggen en wordt onder meer het feit dat de vaccinatie in Nederland zo langzaam van start is gegaan, toegeschreven aan de complexe wirwar van bestuurlijke omwegen van Raden, Platforms en overlegcircuits. In de pers wordt gesproken over schoolbesturen die overheidsgeld incasseren, “dat onvoldoende ten goede komt aan leerlingen en leraren, en al helemaal niet aan het verbeteren van kwaliteit” (Aleid Truijens in de Volkskrant van 15 maart j.l.) en in dezelfde krant sprak Mirjam Janssens over beroepsvergaderars en verandermanagers die “belang hebben bij zoveel mogelijk bestuurlijke

¹⁸ Het aantal onderscheiden stuurmaatregelen is groter dan deze 3 vormen, er is ook nog externe sturing (via de omgeving) en meta- besturing. We laten deze hier buiten beschouwing.

omwegen". Een meer fundamentele analyse over de macht van de bestuurders wordt gegeven door Voermans (2021).

Het is een open vraag of er, ondanks deze aanzwellende kritiek, politiek draagvlak is voor een onafhankelijke evaluatie laat staan een structurele bijstelling van de structuren in het georganiseerde onderwijsveld. Maar ook als dit niet gebeurt heeft de overheid een aantal wettelijke bevoegdheden om impulsen tot verbetering te geven, deze zijn de basis voor het schetsen van enkele mogelijke scenario's in de volgende paragraaf.

6.6 Scenario's

Volgens het meer-niveau model van onderwijseffectiviteit, dat beschreven is in Hoofdstuk 2, is de aanname dat maatregelen op systeemniveau bedoeld zijn om het primaire proces op microniveau in positieve zin te beïnvloeden, met het oog op verhoging van het niveau van prestaties en vermindering van ongelijkheid. In het kader van dit model zijn scenario's voor onderwijsverbetering op te vatten als de doorwerking van stuurmaatregelen op systeemniveau (hefbomen) naar de lagere niveaus.

In de voorafgaande hoofdstukken is al ingegaan op de mogelijkheden en beperkingen van de overheid om de kennis over onderwijseffectiviteit op micro- meso- en macroniveau te gebruiken als hefbomen voor verbetering. Hier vatten we dit samen in drie scenario's. Het eerste scenario lijkt onomstreden en heeft betrekking op het leveren van basisvoorzieningen. Het tweede scenario gaat uit van een meer omvattende analyse en pakt de draad op van het rapport van de Commissie Dijsselbloem uit 2008. Het derde scenario kapitaliseert op examens en toetsing als basis voor curriculum-alignment en incentives om gestructureerd onderwijs te geven.

In deze scenario's komt het terugdringen van ongelijkheid niet expliciet aan de orde. Wel indirect, voor zover er alle aanleiding is om te veronderstellen dat gestructureerd onderwijs in het bijzonder effectief is voor minder bevoorrechte leerlingen. Hetzelfde geldt voor het belang van objectieve toetsing en op uniforme standaarden gebaseerde examens. In hoofdstuk 5 is ongelijkheid in verband met sterk gestratificeerde onderwijsstelsels behandeld. De conclusie was daar dat verbetering van mogelijkheden tot doorstroming tussen schoolcategorieën, zonder het bestel radicaal te hervormen, waarschijnlijk de meest doeltreffende en doelmatige strategie is.

Scenario 1: Opheffing van het lerarentekort; arbeidsvoorzieningen en opleiding voor leraren

Een basisconditie om de kwaliteit van het onderwijs op peil te houden zijn goede arbeidsvoorzieningen voor leraren, met het oog op het terugdringen van het lerarentekort, en een goede initiële lerarenopleiding. Door middel van het onderwijsbudget heeft de overheid hierin een belangrijke functie. In theorie is de kennis over effectieve instructie een belangrijke input voor de lerarenopleiding. De indruk bestaat echter dat de vakdidactische kern minder prioriteit heeft gekregen en dat er erg veel aandacht wordt besteed aan secundaire taken van de leraar als onderzoeker, ontwikkelaar en participant in netwerken. De overheid zou kunnen overwegen om meer invloed uit te oefenen op de inhoud van de curricula voor de lerarenopleiding, en zich hierbij te laten adviseren door een commissie van deskundigen.

Scenario 2: Dijsselbloem 2?

Nogmaals kort samengevat suggereert het overzicht van resultaten van onderwijs-effectiviteitsonderzoek in dit rapport om te kapitaliseren op de kracht van gestructureerd onderwijs en

de ondersteuning daarvan op meso- en macroniveau. Hierbij moet tevens gememoreerd worden dat vakgericht onderwijs in basisvakken de focus is, het gaat om wat soms wordt aangeduid als “basiskwaliteit” en in het rapport van de Commissie Dijsselbloem uit 2008 werd aangeduid als het “kerncurriculum”.

De hefboomen die de overheid ter beschikking staan om het bovenstaande te realiseren hebben te maken met grondwettelijke bevoegdheden op het terrein van het vaststellen van onderwijsdoelstellingen, het curriculum, examens en het toezien op de deugdelijkheid en kwaliteit van het onderwijs. De resultaten van dit rapport zouden kunnen worden opgevat als een aansporing tot het “op punt stellen” van deze bevoegdheden. Er kan niet voorbijgegaan worden aan het historische gegeven dat het precies dit is wat de Commissie Dijsselbloem in 2008 heeft gedaan. De Commissie pleitte onder meer voor leerstandaarden, een duidelijke functie van het Centrale Eindexamen, eventuele uitbreiding van toetsen, verscherpt toezicht door de inspectie een vakgericht kerncurriculum en de eis dat nieuw beleid *evidence based* onderbouwd zou moeten zijn. Ook toen was de opgave het tegengaan van de onderwijsbenadering die hier is aangeduid als de ineffectieve tegenhanger van gestructureerd onderwijs; voor de Commissie Dijsselbloem in 2008 was het op het sociaal constructivisme gebaseerde “nieuwe leren” en voor Duijker was in het in 1976 het “open” onderwijs dat in de Contourennota werd bepleit. Anno 2021 is precies hetzelfde syndroom terug met als meest actuele label het gepersonaliseerde leren. Deze historische reconstructie roept vragen op, die de reikwijdte van deze studie te buiten gaan. Er zou een soort nieuwe commissie Dijsselbloem nodig zijn om de tentatieve diagnose te verifiëren, dat de aanbevelingen uit 2008 op kernpunten niet zijn nageleefd en dat niet effectieve vormen van onderwijs opnieuw een te grote invloed hebben. Daarbij zou tevens aandacht kunnen worden besteed aan de achtergronden hiervan, zowel wat betreft alternatieve visies, achterliggende theorie, als de rol van het georganiseerde middenveld van het onderwijsstelsel.

Scenario 3: Examenvoorbereiding en uitlijning van het curriculum

De grote autonomie van de scholen en de lumpsum financiering betekenen dat de overheid geen rechtstreeks invloed heeft op het gebruik van de kennis over instructie-effectiviteit. Een voorlichtingscampagne zou wel tot de mogelijkheden behoren. Ook heeft de overheid indirect invloed via de contouren van het curriculum, de doelstellingen en eindtermen en de examens. Ook het centrale examen en de eindtoets zijn instrumenten voor de overheid om invloed uit te oefenen op de inhoud van het onderwijs. Volgens de theorie over curriculum-alignment zullen scholen en methoden ontwikkelaars zich richten op het examen; de kwaliteit van het examen is tevens de belangrijkste conditie om de marktwerking gefocust te houden op de doelstellingen. Ook verplichte leerlingvolgsystemen kunnen aan de goede uitlijning van het onderwijs bijdragen. In een stelsel als het Nederlandse, met een hoge mate van decentralisatie en schoolautonomie, vervult het eindexamen een cruciale functie voor het in stand houden van de kwaliteit van het onderwijs. Het examen regelt het civiel effect van het onderwijs voor leerlingen, stimuleert en motiveert leerlingen tot prestatiegerichtheid, biedt richting aan de ontwikkeling van leermethoden en de doelstellingsgerichtheid van het onderwijs en is, *last but not least*, van groot belang om scholen en schoolbesturen verantwoordelijk te houden voor de prestaties ('accountability').

6.7 Suggesties voor vervolgactiviteiten

Follow-up van deze studie

Na goedkeuring van dit rapport door de opdrachtgever zou overwogen kunnen worden om over te gaan tot bredere verspreiding en disseminatie; eventueel in de vorm van een aangepaste, op de onderwijspraktijk gerichte versie van het rapport.

Uitbreiding van de literatuur inventarisatie over meta-analyses na 2015. Een aanzet is gegeven in Annex 2, waarin de focus lag op internationaal vergelijkend onderzoek.

Nieuw fundamenteel wetenschappelijk onderzoek dat gesuggereerd is in deze studie

Uit het overzicht in de Annexen 1 en 2 blijkt dat de data van PISA en TIMSS een rijke bron zijn voor nadere analyses. Vooral voor TIMSS geldt dat er nog veel interessant materiaal “braak ligt” en dat de bestanden eigenlijk “ondergeanalyseerd” zijn.

Een permanente service voor het uitvoeren van secundaire analyses op resultaten van TIMSS en PISA zou een goede investering zijn.

In Annex 3 is een eerste aanzet gegeven voor nader onderzoek naar de mogelijkheden van onderwijstechnologie en het gebruik van artificiële intelligente (AI) in het kader van instructie-effectiviteit.

Onderzoek naar en ontwikkeling van directe instructie op het terrein van hogere cognitieve vaardigheden en meta-cognitie. Hierbij kan worden aangesloten bij het promotieonderzoek van Marinka Drost en Petra Verra, begeleid door Bernard Veldkamp en Jaap Scheerens en bijdragen van de RUG.

Een representatief onderzoek naar onderwijsconcepten

Research en ontwikkeling op het terrein van kwaliteitsmonitoring

Het onder de aandacht brengen van bestaande systemen om de kwaliteit van toetsen en examens te beoordelen en het in opdracht uitvoeren van beoordelingen van toetsen en examens.

Inventarisatie en kwaliteitsbeoordeling van op toetsprogramma's en volgsystemen gebaseerde programma's voor schoolzelfevaluatie

Voortgezet onderzoek naar de meetbaarheid van sociaal-emotionele vaardigheden (voortgezet tov soft skills review door Scheerens, Van der Werf en De Boer, 2020, en onderzoek door de UVA in kader van Nationaal Cohort Onderzoek)

Herontwerp van kwaliteitskaarten van schoolprestatie-indicatoren.

Onafhankelijk beleidsgericht evaluatieonderzoek (ex ante zowel als ex post)

De studie levert verder thema's voor onafhankelijk beleidsgericht evaluatieonderzoek op (liefst vanuit een productiviteitsperspectief, waarbij naast opbrengsten ook kosten worden vastgesteld). Mogelijke thema's:

- het functioneren van lerarenopleidingen;
- onderzoek naar kosten en baten van bestuurlijke overhead;
- kosten-baten analyse van praktijkgericht onderwijsonderzoek.

Bijlage 1 Literatuur Referenties

- Allison, G.T., (1969) Conceptual models and the Cuban missile crisis. *American Political Science Review*. 63 3.
- Allison, G.T., (1971) *Essence of Decision*. Boston: Little, Brown & Co.
- Aloisi, C., and Tymms, P., (2017) PISA trends, social changes, and education reforms. *Educational Research and Evaluation*. 23 180- 220.
- Amelvoort, H.W.C.H. van, & Scheerens, J. (1997). Policy issues surrounding processes of centralization and decentralization in European education systems. *Educational Research and Evaluation*, 3(4), 340-363.
- Anderson, L.W., (2004) *Increasing Teacher Effectiveness*. (Second edition) Paris: IIEP, UNESCO
- Argyris, C., & Schön, D.A. (1974). *Theory in practice: increasing professional effectiveness*. San Francisco: Jossey-Bass.
- Artelt, C. Baumert, J., McElvany, N., and Peschar, J.) (2003) *Learners for life. Student approaches to learning. Results from PISA 2000*. Paris: OECD publishing.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. C. (1983). Effects of coaching program on achievement test performance. *Review of Educational Research*, 53, 571-586.
- Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213– 238.
- Baumert, J., Blum, W., Brunner, M., Jordan, A., Klusmann, U., Krauss, S., Kunter, M., Neubrand, M., Yi-Miau Tsai (2005) Professional knowledge of teachers, cognitively activating instruction, and the development of mathematical competence (COACTIV). Presentation at OECD Task Force on Teaching Effectiveness. Reykjavik, September, 2005.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133-180.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom.
- Bereiter, C., & Kurland, M. (1996). A constructive look at Follow Through results. *Effective School Practices*, 15, 17–32.

- Biesta, G. (2017) Lead Learner or Head Teacher? Exploring Connections Between Curriculum, Leadership and Evaluation in an 'Age of Measurement'. In: Uljens, M., and Ylimkai, R. M. (2017) *Bridging the gap between Educational Leadership, Curriculum Theory and Didaktik. Non-affirmative theory of education*. Springer Open (180- 198).
- Biesta, G. (2018) Tijd voor pedagogiek. Over de pedagogische paragraaf in het onderwijs. (*Time for pedagogics. About the pedagogical paragraph in education*). Utrecht: Universiteit voor Humanistiek. Inaugural Lecture.
- Bishop, J. (1997). *The effect of National Standards and Curriculum-Based Exams on Achievement*. Cornell University. Center for Advanced Human Resource Studies.
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.
- Blank, J. L. T., & van Heezik, A. A. S. (2015). *Productiviteit van overheidsbeleid, deel I: het Nederlandse onderwijs, 1980-2012*. Den Haag/Delft: Eburon.
- Blank, J.L.T.< & van Heezik, A.A.S. (2019) Productiviteit van de Overheid. IPSE studies. Delft: IPSE file:///C:/Users/Jaap/AppData/Local/Temp/IPSE1801-3_Productiviteitoverheid.pdf
- Blömeke, S., Busse, A., Kaiser, G., Koenig, J. and Suhl, U. (2016) The relation between content-specific and general teacher knowledge and skills. *Teaching and Teacher Education*. 56, 35-46
- Bloom, B.S., Hastings, J.Th., & Madaus, G.F. (1971). *Handbook on Formative and Summative Evalaution of Student Learning*. New-York: MacGraw-Hill.
- Boer, de, H., Donker, A.S., Kostons, D.N.M. & Van der Werf, G.P.C., (2018) Long-Term Effects of Metacognitive Strategy Instruction on Student Academic Performance: A Meta-Analysis. *Educational Research Review* 24, 98- 215.
- Bol, T., Wischge, J., Van de Werfhorst, H.G. & Dronkers, J. (2014) Curricular Tracking and Central Examinations: Counterbalancing the Impact of Social Background on Student Achievement in 36 countries. *Social Forces* 92(4) 1545–1572.
- Bosker, R.J., & Scheerens, J. (1999). Openbare prestatiegegevens van scholen; nuttigheid en validiteit. *Pedagogische Studiën*, 76(1), 61-73.
- Bronneman-Helmerts, R, and the Social Cultural Planning Bureau (2011), *Overheid en onderwijsbestel Beleidsvorming rond het Nederlandse onderwijsstelsel, (1990-2010)* The Hague:, SCP-publication 2011-31.
- Brophy, J., & Good, T.L. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Third Handbook of Research on Teaching* [pp. 328-375]. New York: MacMillan.
- Brunello, G. And Checchi, D., (2006) Does School Tracking Affect Equality of Opportunity? IZA Working Paper. Bonn.

- Bryk, T., Thum, Y., Easton, J., & Luppescu, S. (1998). *Charting Chicago school reform*. Boulder, CO: Westview Press.
- Carnoy, M., Elmore, R., & Siskin, L. (2003) (Eds.). *The New Accountability. High Schools and High-Stakes Testing*. New York/London: Routledge Falmer.
- Carroll, J.B. (1989) The Carroll model, a 25-year retrospective and prospective view. *Educational Researcher*, 16, 26 -31.
- Causa, O., & Chapuis, C. (2009). *Equity in Student Achievement Across OECD Countries: An investigation of the role of policies*. OECD Economics Department Working Papers, No. 708. OECD Publishing.
- Chibulka, J.G. & Derlin, R.L. (1995). State educational performance reporting policies in the U.S.: accountability's many faces. *International Journal of Educational Research*, 23(6), 479-492.
- Chubb, J.E., & Moe, T.M. (1990). *Politics, markets and American schools*. Washington, D.C.: Brookings Institute.
- Coelli, M., & Green, D.A., (2012) Leadership effects: School principals and student outcomes. *Economics of Education Review*, 31, 1, 92- 109.
- “Commissie Dijsselbloem” Parlementair onderzoek Onderwijsvernieuwingen (2018) Tweede Kamer, vergaderjaar 2007–2008, 31 007, nr. 6
- Connolly, M., & Kruse, S.D., (2019). Organizational culture in schools: A review of a widely misunderstood concept. In : Connolly, M., Eddy-Spicer, D.H., James, C. and Kruse, S.D. (eds.) *The Sage Handbook of school organization*. Sage Publishing Ltd.
- Daily, A.J., Moolenaar, N.M. and Carrier, N. (2011) Reform at the edge of chaos: connecting complexity, social networks, and policy implementation. (manuscript)
- De Corte E. et al. (1973). *Beknopte didaxologie*. Groningen: H.D. Tjeenk Willink.
- Darling-Hammond, L. (1999). *Teacher Quality and Student Achievement: A Review of State Policy Evidence*. Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.
- Dietrichson, J., Boeg, M., Fliges, T., & Klint-Joergensen, A. (2017) Academic Interventions for Elementary and Middle School Students With Low Socioeconomic Status: A Systematic Review and Meta-Analysis. *Review of Educational Research* 87(2): 243-282.
- Donker, A.S. (2015) Towards effective learning strategies. Proefschrift. Rijks Universiteit Groningen.
- Donley, J. and States, J. (2020) School principal competencies. Wing Institute Original Paper https://www.academia.edu/41536601/School_Principal_Competerencies_Wing_Institute_Original_Paper_email_work_card=view-paper

- Doyle, W. (1985). Effective secondary classroom practices. In: M.J. Kyle (ed.), *Reaching for excellence. An effective schools sourcebook*. Washington DC: US Government Printing Office.
- Driessen, (2018) Zo doorgaan met het onderwijsachterstandenbeleid? Didactief online 04-01-2018 <https://didactiefonline.nl/blog/blonz/zo-doorgaan-met-het-onderwijsachterstandenbeleid>
- Drost, M., & Verra, P. (2019). *Handboek RTTI met theoretische beschouwing*. (Handbook RTTI, with a theoretical reflection). Bodegraven, The Netherlands: Docentplus.
- Duiker, H.C.J. (1976) De ideologie van de zelfontplooiing. (The ideology of self-development). *Pedagogische Studieren*. (53) 358-373.
- Effrem, K., Robbins, J., & Pioneer Institute for Public Policy Research. (2019). Social-Emotional Learning: K-12 Education as New Age Nanny State. White Paper No. 192. In *Pioneer Institute for Public Policy Research*. Pioneer Institute for Public Policy Research.
- Elmore, R.F. and Associates (1990) *Restructuring School; The Next Generation of Educational Reform*. San Francisco: Jossey Bass.
- Fend, H. (2006). *Neue Theorie der Schule*. VS Verlag für Sozialwissenschaften.
- Fend, H. (1980). *Theorie der Schule*. München: Urban & Schwarzenberg.
- Fuchs, L.S. & Fuchs, D. (1985). *A Quantitative Synthesis of Effects of Formative Evaluation on Achievement*. [Washington, D.C.]: Distributed by ERIC Bosker, R.J., & Scheerens, J. (1999). Openbare prestatiegegevens van scholen; nuttigheid en validiteit. *Pedagogische Studiën*, 76(1), 61-73.
- Fuchs, Th., & Wößmann, L. (2004). *What accounts for international differences in student performance? A re-examination using PISA data*. CESifo Working paper No. 1235. University of Munich, Munich, Germany.
- Fullan, M., & Watson, N. (2000). School-Based-Management: Reconceptualizing to Improve Learning Outcomes. *School Effectiveness and School Improvement*, 11(4), 453-473.
- Gage, N. (2009) *A conception of teaching*. Springer Science + Business Media
- Glass, G.V., & Matthews, D.A. (1991). Are Data Enough? Review of Chubb and Moe's "Politics, Markets and America's Schools". *Educational Researcher*, 29(3), 24-27.
- Guthrie, G. (2021) *Foundations of classroom change in developing countries. Part 1 Evidence* Available from Research Gate Download.
- Guthrie, G. (2011) The Progressive Education Fallacy in Developing Countries. In favor of formalism. Springer. <https://www.springer.com/gp/book/9789400718500>
- Good, Wiley & Florez (2009). Effective Teaching: An Emerging Synthesis. In G. Dworkin (Eds.), *International Handbook of Research on Teachers and Teaching*. New York: Springer, pp. 803-816.

- Hanushek, E., and Rivkin, S.G. (2010) Generalizations about using value-added measures of teacher quality *American Economic Review: Papers & Proceedings* 100: 267–271.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81–112.
- Hattie, J. (2009) *Visible Learning*. New York, London: Routledge.
- Heck, R. H. & Moriyama, K.(2010). Examining relationships among elementary schools' contexts, leadership, instructional practices, and added-year outcomes: a regression discontinuity approach. *School Effectiveness and School Improvement*, 21(4), 377-408.
- Hendriks, M.A., and Scheerens, J.. (2014) Effects of evaluation and assessment on student achievement: A review and meta-analysis. Enschede: University of Twente.
- Hendriks, M.A. and Scheerens, J. (2013) School leadership effects revisited: a review of empirical studies guided by indirect-effect models. *School leadership & management*, 33 (4). pp. 373-394.
- Hendriks, M.A. (2014) *The influence on school size, leadership, evaluation, and time on student outcomes*. Enschede: University of Twente, doctoral thesis.
- Holcomb, R., Jennings, J., & Koretz, D. (2013) The roots of score inflation: An examination of opportunities in two states' tests. In: G. Sunderman (ed.) *Charting reform, achieving equity in a diverse nation* (163-189). Greenwich, CT: Information Age Publishing.
- Honingh, M.E., and Ehren, M.C.M. (2012) *Bestuurskunde*, vol 21 (4) 64 – 72.
- Jimenez, E., & Sawada, Y. (1998). *Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program* (Working Papers Series on Impact Evaluation of Education Reforms No. 8). Washington, DC: World Bank.
- Kickert, (1979) *Organization of decision making. A systems-theoretical approach*. Amsterdam: North Holland Publishing Company.
- Kim, S-E. (2005). Effects of implementing performance assessments on student learning: meta-analysis using HLM. (Unpublished Doctoral dissertation). University Park, PA: The Pennsylvania State University.
- Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30(4), 28–37.
- Kirschner, P.A., Claessens, L. & Raaijmakers, S. (2019) *Op de schouders van reuzen Inspirerende inzichten uit de cognitieve psychologie voor leerkrachten (Standing on the shoulders of giants. Inspiring insights from cognitive psychology for teachers*.
<https://newsroom.didactiefonline.nl/uploads/BOEKEN/20190107%20Op%20de%20schouders%20van%20reuzen%20Definitief%20download.pdf>

- Kirschner, P. (2020) Progressief slechter. Didactief -online, 01-09-2020
<https://didactiefonline.nl/blog/paul-kirschner/progressief-slechter>
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41, 75-86.
- Klieme, E., (2012) Qualities and effects of teaching. Towards a conceptual theory of teaching. Key note address. EARLI SIG meeting, Educational Effectiveness, Zuerich, 23 August.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284.
- Kools, M. and Stoll L. (2016), “What Makes a School a Learning Organisation?”, OECD Education Working Papers, No.137, Paris: OECD Publishing.
- Koretz, D. (2005) Alignment, High Stakes, and the Inflation of Test Scores. *Yearbook of the National Society for the study of Education*. 104, 2 99-118.
- Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students’ academic, behavioral emotional and motivational outcomes. *Review of Educational Research*. 86 (3) 643 – 680.
- Kramer, N.J.T.A. (1978) *Systeem in probleem. Een onderzoek naar de bijdragen van de systeemleer tot de aanpak van praktijkproblemen in organisaties*. Leiden: Stenfert Kroese.
- Krueger, M. & Scheerens, J. Conceptual perspectives on school leadership. In Scheerens, J. (Ed.). (2012) *School leadership effects revisited: Review and meta-analysis of empirical studies*. Dordrecht: Springer (pp 1- 30).
- Kuijk, J. van, Van Langen, A., Driessen, G., & Elfering S. (2015) *Procesevaluatie en effectmeting programma School aan Zet*. Nijmegen, ITS.
- Kurz, A., Elliott, S.N., Wehby, J.H. and Smithson, J.L. (2010) Alignment of the Intended, Planned, and Enacted Curriculum in General and Special Education and Its Relation to Student Achievement. *The Journal of Special Education*, 44(3) 131–145.
- Kurz, A. (2011) *Opportunity to learn the intended curriculum: measuring key instructional indicators and examining relations to achievement rot students with disabilities*. Doctoral dissertation. Vanderbilt University. Nashville Tennessee.
- Kyriakides, L., Creemers, B., Antoniou, P., & Demetriou, D. (2010). A synthesis of studies searching for school factors: implications for theory and research. *British Educational Research Journal*, 36(5), 807-830.

- Kyriakides, L., and Luyten, H. (2009) The contribution of schooling to the cognitive development of secondary education students in Cyprus: an application of regression discontinuity with multiple cut-off points. *School Effectiveness and School Improvement*, 20(2), 167-186.
- Kyriakides, L., Christoforou, C. & Charalambous, C.L. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education*, 36(November), 143-152.
- Lee, S-Y. (2007). *Structural equation modelling: A Bayesian approach*. Chichester, England: Wiley.
- Leeuw, de, A.C.J. *Organisaties, management, analyse, ontwerp en verandering. Een systeemvisie*. Assen: Van Gorkum
- Leithwood, K., & Menzies, T. (1998b). A Review of Research Concerning the Implementation of Site-Based Management. *School Effectiveness and School Improvement*, 9(3), 233-285.
- Leithwood, K., & Menzies, T. (1998a). Forms and Effects of School-Based Management: A Review. *Educational Policy*, 12(3), 325-246.
- Leithwood, K.A., & Montgomery, D.J. (1982). The role of the elementary school principal in program improvement. *Review of Educational Research*, 52, 309-399.
- Levačić, R. (1995). *Local management of schools: Analysis and practice*. Buckingham: Open University Press.
- Liebowitz, D.D., and Porter, L. (2020) The Effect of Principal Behaviors on Student, Teacher, and School Outcomes: A Systematic Review and Meta-Analysis of the Empirical Literature. *Review of Educational Research* 89(5), 785 –827.
- Lomos, C., Hofman, R.H. & Bosker, R.J. (2011) Professional communities and student achievement – a meta-analysis, *School Effectiveness and School Improvement*, 22:2, 121-148.
- Lomos, C., Hofman, R.H., & Bosker, R.J. (2011a) The relationship between departments as professional communities and student achievement in secondary *Teaching and teacher education* 27 (4) 2011, 722-731.
- Lotto, I., & Clark, D.L. (1986) Understanding planning in educational organizations. *Planning and Change*, 19, 9-18.
- Luyten, H, and Scheerens, J. (2014) Analysis of change in national reading literacy averages: correlations with changes in system level and school characteristics. Manuscript. Enschede: University of Twente.
- Luyten, J. W., Scheerens, J., Visscher, A. J., Maslowski, R., Witziers, B. U., & Steen, R. (2005). *School factors related to quality and equity. Results from PISA 2000*. Paris: OECD.

- Malen, B., Ogawa, R.T., & Kranz, J. (1990). What do we know about school-based management?: a case study of the literature – a call for research. In W.H. Clune & J.F. Witte (Eds.), *Choice and Control in American Education (Vol. 2): The Practice of Choice, Decentralization and School Restructuring* [pp. 289-342]. London: Falmer Press.
- Marzano (2003) *What works in schools. Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Maslowski, R. (2001). *School Culture and School Performance*. An explorative study into the organizational culture of secondary schools and their effects. Enschede: Twente University Press (dissertation).
- Masuch, M., LaPotin, P. and Verhörst, P. 'Drowning in a sea of chewing gum'. Makework among members of bureaucratic organizations. *Amsterdams Sociologisch Tijdschrift*, jrg. 15, nr. 4, maart 1989.
- McEwan, P.J. (2000). *Comparing the Effectiveness of Public and Private Schools: A Review of Evidence and Interpretations* (National Center for the Study of Privatization in Education Occasional Paper No. 3). New York, NY: NCSPE, Columbia University.
- McKinsey & Company (2020) Een versterkt fundament voor iedereen. Een onderzoek naar de doelmatigheid en toereikendheid van het funderen onderwijs (primair en voortgezet). McKinsey & Company, the Netherlands.
- McKinsey & Company -Mourshed, M., Chijioke, C. & Barber, M. (2010). *How the world's most improved school systems keep getting better*.
- McMeekin, R.W. (2003). *Incentives to improve education: a new perspective*. Cheltenham: Edward Elgar.
- Meelissen, M., Hamhuis, E & Weijn, L. (2020) Leerlingprestaties in de exacte vakken in groep 6 van het basisonderwijs. Resultaten TIMSS-2019. Universiteit Twente.
- Messick, S. (1982). Issues of effectiveness and equity in the coaching controversy: Implications for educational and testing practice. *Educational Psychologist*, 17, 67–91.
- Messner, R., & Blum, W. (2019). Der Mythos des offenen Unterrichts – unter Einbeziehung von Befunden aus dem DISUM-Projekt. In: Steffens & Messner (Eds.) *Unterrichtsqualitaet. Konzepte und Bilanzen gelingenden Lehrens and Lernens*. Muenster- New York: Waxmann.
- Meyer, J.W., and Rowan, B. (1977). "Institutionalized Organizations: Formal Structure as Myth and Ceremony," *American Journal of Sociology*, 83: 340-63.
- Meyer, J.W., & Rowan, B. (1992) Institutionalized organizations: Formal structure as myth and ceremony. In J.W. Meyer and W.R. Scott (Eds.) *Organizational environments: ritual and rationality*. London: Sage.
- Ministerie van OCW (2007). *Scholen voor morgen*. Kwaliteitsagenda PO. Den Haag: Ministerie OCW.

- Ministerie van OCW (2007a). *Tekenen voor Kwaliteit*. Kwaliteitsagenda Voortgezet Onderwijs. Den Haag: Ministerie OCW.
- Ministerie van OC&W (2013) *Lerarenagenda 2013-2020: de leraar maakt het verschil*. Den Haag: Ministerie OCW.
- Mintzberg, H. (1979). *The structuring of organizations*. Englewood Cliffs, NJ: Prentice-Hall.
- Monk, D.H. (1992). *Microeconomics of School Productions*. Paper for the Economics of Education Section of the International Encyclopedia of Education.
- Monk, D.H. (1989). The education production function: it's evolving role in policy analysis. *Educational Evaluation and Policy Analysis*, 11(1), 31-45.
- Morgan, G. (1986). *Images of organizations*. Beverly Hills: Sage.
- Mulder, L., en Meijnen, W., (2013) *Onderwijsachterstanden in de BOPO-period, 2009- 2012. Een Review*. ITS Nijmegen.
- National Educational Policy Center NEPC (2019) *Personalized Learning and the Digital Privatization of Curriculum and Teaching*
<https://nepc.colorado.edu/publication/personalized-learning>
- Niskanen, W.A., (1971) *Bureaucracy and representative government*. Chicago: Aldine-Atherton.
- OECD, (2010) *PISA 2009 results: What makes a school successful? Resources, policies and practices*. Volume IV. Paris: OECD Publishing.
- OECD (2010a) *Strong Performers and Successful Reformers in Education: Lessons from PISA for the United States*. Paris: OECD.
- OECD (2013) *PISA 2012 Results, Volume IV, What makes schools successful? Resources, policies and practices*. Paris: OECD Publishing.
- OECD (2014) *OECD Reviews of evaluation and assessment in education; Netherlands 2014*. Paris: OECD publishing.
- OECD (2016) *PISA 2015 Results. Policies and practices for successful schools. Volume II*. Paris: OECD Publishing.
- OECD (2019) *PISA 2018 Results. Effective policies, successful schools*. Volume V. Paris: OECD Publishing.
- Onderwijsraad (2015) *Maatwerk binnen wettelijke kaders: eindtoetsing als ijkpunt voor het funderend onderwijs*. Den Haag: Onderwijsraad.
- Oomens, M., Exalto, R., De Jong, A., Scholten, F., Veldkamp B., Janse, R., & Scheerens, J. (2017) *Marktonderzoek formatief evalueren. Een onderzoek naar vraag en aanbod*. Utrecht: Oberon.

- Peal, R. (2014). Progressively worse: The burden of bad ideas in British schools. London, VK: Civitas
<http://www.civitas.org.uk/pdf/ProgressivelyWorsePeal.pdf>
- Polikoff, M.S. and Porter, A.C. (2014) Instructional alignment as a measure of teaching quality.
Educational Evaluation and Policy Analysis, 36, 399-416.
- Popham, W.J. (2003) *Test better, teach better*, The instructional role of assessment Alexandria, Virginia: ACSD.
- Praetorius, A. K., Klieme, E., Kleickmann, T., Brunner, E., Lindmeier, A., Taut, S., & Charalambous, C. (2020). Towards developing a theory of generic teaching quality: Origin, current status, and necessary next steps regarding the Three Basic Dimensionsmodel. *Zeitschrift für Pädagogik*, Beiheft 66, 15-36.
- Purkey, S.C., & Smith, M.S. (1983). Effective schools: a review. *The Elementary School Journal*, 83(4), 427-452.
- Quinn, R.E. and Rohrbauch 1983 Spatial models of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science*, 29, 363 – 377.
- Rand News Release (July 25, 2000). Explaining achievement gains in North Carolina and Texas.
- Ravens, J. van (2011) The selectivity and vocational orientation of education systems. In Scheerens, J., Luyten, H., and Van Ravens, J. (2011). *Perspectives on educational quality. Illustrative outcomes on primary and secondary schooling in the Netherlands*. Dordrecht, Heidelberg, New-York, London: Springer.
- Reynolds et al. (2014) Reynolds, D., Sammons, P., De Fraine, B., Townsend, T., Van Damme, J., Teddlie, C., &Stringfield, S. (2012). Educational Effectiveness Research (EER): A state of the Art Review: *School Effectiveness and School Improvement*, 25 197 - 230.
- Scheerens, J. (2017a) *Ed. Opportunity to learn, curriculum alingment and test preparation. A research review*. Dordrecht, New York, London: Springer.
- Scheerens, J. (1992). *Effective Schooling, Research, Theory and Practice*. London: Cassell.
- Scheerens, J. & Bosker, R. (1997) *The Foundations of Educational Effectiveness*. Oxford: Pergamon
- Scheerens, J., Luyten, H., Steen, R. Luyten-de Thouars, Y. (2007). *Review and Meta-Analyses of School and Teaching Effectiveness*. Enschede: University of Twente, Department of Educational Organization and Management.
- Scheerens, J. (2008) L'istruzione Intesa come Sistema adattivo complesso: implicazioni per gli studi sull'efficacia educative. In Abbona, F., Del Re G., & Monaco, G. *Complessita dinamica dei processi educative. Aspetti teorici e pratici*.

- Scheerens, J., Luyten, H., and Van Ravens, J. (2011). *Perspectives on educational quality. Illustrative outcomes on primary and secondary schooling in the Netherlands*. Dordrecht, Heidelberg, New-York, London: Springer.
- Scheerens, J. (Ed.). (2012) *School leadership effects revisited: Review and meta-analysis of empirical studies*. Dordrecht: Springer.
- Scheerens, J. (2013) *What is effective schooling? A review of current thought and practice*. International Baccalaureate Organization.
file:///C:/Users/Jaap/AppData/Local/Temp/WhatisEffectiveSchoolingFINAL.pdf
- Scheerens, J., & Doolaard, S. (2013) Review studie Onderwijskwaliteit PO. Groningen: GION.
- Scheerens, J., (2014) School, teaching, and system effectiveness: some comments on three state-of-the-art reviews. *School effectiveness and school improvement*, 25 (2). 282 – 290.
- Scheerens, J., Luyten, H., Glas, C.A., Jehangir, K., and Van den Bergh, M. (2014) System level indicators. Analyses based on PISA 2009 data. Internal report. Enschede: University of Twente.
- Scheerens, J. Luyten, H., Berg, S. M. van den and Glas, C. A.W. (2015) Exploration of direct and indirect associations of system-level policy-amenable variables with reading literacy performance. *Educational research and evaluation*, 21 (1). 15 – 39.
- Scheerens, J. (2016) *Educational effectiveness and ineffectiveness: A critical review of the knowledgebase*. Dordrecht: Springer.
- Scheerens, J. & Exalto, R., (2017) *“Teaching to and from the test”. Een verkennende studie naar het in kaart brengen van doelen, toetsen, curriculum en onderwijsaanbod*. Utrecht: Oberon.
- Scheerens, J. (2017). The perspective of limited malleability in educational effectiveness: treatment effects in schooling. *Educational Research and Evaluation*. 23 247- 266.
- Scheerens, J. (2017a) *Ed. Opportunity to learn, curriculum alingment and test preparation. A research review*. Dordrecht, New York, London: Springer.
- Scheerens, J. Brouwer, A., Sanders, P., Veldkamp, B, en Van der Vegt, A. (2019) Fundamentele vragen over examens. Utrecht: Oberon.
- Scheerens, J., Timmermans, A., & Van der Werf, G. (2019) Socioeconomic inequality and student outcomes in the Netherlands. In: Volantes, L. Schnepf, J.J. and Klinger, D.A. (Editors). *Socioeconomic inequality and student outcomes. Cross-national trends, policies, and practices*. Singapore: Springer Nature.
- Scheerens, J., Van der Werf, G., & De Boer, H. (2020) *Soft skills in education. Putting the evidence in perspective*. Cham, Switzerland: Springer Nature.

- Scheerens, J. (2021) Theories on teaching effectiveness. Manuscript under review.
- Schein, E. H. (1985). *Organizational Culture and Leadership: A Dynamic View*. San Francisco, CA: Jossey-Bass.
- Schleicher, A. (2018), *Valuing our Teachers and Raising their Status: How Communities Can Help*, International Summit on the Teaching Profession, OECD Publishing, Paris.
- Schroeder, C.M., Scott, T., Tolson, H., Tse-Yang Huang, Yi-Hsuan Lee (2007) A Meta-Analysis of National Research: Effects of Teaching Strategies on Student Achievement in Science in the United States. *Journal of research in science teaching*. 44, (10), 1436–1460.
- Schreyer, P., Herbert, B., Lindmeier, A. (2020) Multiperspektivische Betrachtung der kognitiven Aktivierung mit Beispielen aus der TALIS-Videostudie. | ZIB-Fachtagung 03.12.2020 | Workshop Multiperspektivische Betrachtung der kognitiven Aktivierung. Frankfurt: DIPF.
- Seidel, T., And Steen, R. (2005) The indicators on teaching and learning compared to the review of recent research articles on school and instructional effectiveness. In: Scheerens, J., Seidel, T.,
- Seidel and Shavelson, 2007 Seidel, T., & Shavelson, R.J. (2007). Teaching effectiveness research in the past decade: the role of theory and research design in disentangling meta-analysis results. *Review of Educational Research* 77(4), 454-499.
- Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4–14.
- Slavin, R.E. (1996). *Success for all*. Lisse: Swets & Zeitlinger.
- Somers, M-A., McEwan, P.J., & Willms, J.D. (2004). How effective are private schools in Latin-America. *Comparative Education Review*, 48, 48-69.
- Spada, N.& Tomita, Y. (2010) Interactions Between Type of Instruction and Type of Language Feature: A Meta-Analysis. *Language Learning* 60:2, 263–308.
- Stacey, R.D., Griffin, D., & Shaw, P. (2000) *Complexity and management. Fad or radical challenge to systems thinking?* New York: Routledge.
- Stallings, J., & Mohlman, G. (1981). *School policy, leadership style, teacher change and student behavior in eight schools*. Final report to the National Institute of Education, Washington D.C.
- Stockard, J., Wood, T.W., Coughlin, C., and Khoury, C.R (2018) The Effectiveness of Direct Instruction Curricula: A Meta-Analysis of a Half Century of Research. *Review of Educational Research* Vol. 88, No. 4, pp. 479 –507.
- Stringfield, S. (1995). Attempting to enhance students' learning through innovative programs: the case for schools evolving into high reliability organizations. *School Effectiveness and School Improvement*, 6(1), 67-96.

- Toole, J. C., & Louis, K. S. (2002). The role of professional learning communities in international education. In K. Leithwood, & P. Hallinger (Eds.), *Second international handbook of educational leadership and administration* (pp. 245e279). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Van der Ley, D.A.V. (2020) *Hoe kunnen PO-scholen er voor zorgen dat het leesonderwijs op orde komt?* Research Ed 202
<https://files.sgbsg.nl/redeu/uploads/2020/04/21141816/Leij-researchEd-2020.pdf>
- Veldkamp, B. (2019) Innovaties in toetsen en examens in het VO. In; Scheerens, J. Brouwer, A., Sanders, P., Veldkamp, B, en Van der Vegt, A. (2019) *Fundamentele vragen over examens*. Utrecht: Oberon.
- VH educatieve sector (2019) *Strategische beleidsagenda 2019- 2024*. Samen toekomstbestendige leraren opleiden.
https://www.vereniginghogescholen.nl/system/knowledge_base/attachments/files/000/001/088/original/085_040_SBA_EDUCATIE_BW_DEFDEF.pdf?1575577347
- Visscher, A. J., & Ehren, M. (2011). *De eenvoud en complexiteit van opbrengstgericht werken*. [Enschede]: [Universiteit Twente, Vakgroep Onderwijsorganisatie en -management].
- Visscher, A.J. (2020) On the value of data-based decision making in education: The evidence from six intervention studies. *Studies in Educational Evaluation*,
<https://doi.org/10.1016/j.stueduc.2020.10089>
- Vijfeijken, M. van, en anderen (2010) *Meta-analyse van de eerste opbrengsten van de Kwaliteitsagenda Primair Onderwijs IVA en ITS, 2010*.
- Vos, H. de (1989). A rational-choice explanation of composition effects in educational research. *Rationality and Society*, 1(2), 220-239.
- Walber, H.j., Paik, S.J., Komukai, A. & Freeman, K. (2000). Decentralization: An international perspective. *Horizons*, Winter, 2000.
- Warner, W.K. and Havens, A.E. (1968) Goal displacement and the intangibility of organizational goals. *Administrative Science Quarterly*. 12 (4) 539 – 555.
- Wayne, A. J., & Youngs, P. (2003). Teacher Characteristics and Student Achievement Gains: A Review. *Review of Educational Research*, 73, 89-122.
- Weick, K.E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1-19.
- Werf, M. P. C. van der (2005). *Leren in het studiehuis: consumeren, construeren of engageren? (Learning in the 'Study House', consuming, constructing or engaging?)* Groningen: Rijksuniversiteit Groningen.
- Werfhorst, H. van de, J. Mijs (2010), "Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective", *Annual Review of Sociology*, Vol.36/1, pp.407-428, <http://dx.doi.org/10.1146/annurev.soc.012809.102538>

- Winkler, D.R., & Gershberg, A.I. (2000). *Education Decentralization in Latin America: The Effects on the Quality of Schooling* (LCSHD Paper Series No. 59). Washington, DC: World Bank.
- Whitty, G., Power, S., & Halpin, D. (1998). *Devolution and Choice in Education: The School, the State and the Market*. Buckingham: Open University Press.
- Williamson, B. (2019). Psychodata: disassembling the psychological, economic, and statistical infrastructure of 'social-emotional learning'. *Journal of Education Policy*.
- Willms, J.D., & Somers, M.-A. (2000). *Schooling outcomes in Latin America*. Report prepared for UNESCO-OREALC and the Laboratorio Latinoamericano de la Calidad de la Educación [The Latin American Laboratory for the Quality of Education].
- Witziers, B., Hendriks, M., & Doornekamp, G. (2005). *Positioning the supervision frameworks for primary and secondary education of the Dutch Educational Inspectorate in current educational discourse and validating core indicators against the knowledge base of educational effectiveness research*. Enschede/Kiel: University of Twente / Institute for Science Education (IPN).
- Wößmann, L. (2018) Central exit exams improve student outcomes. External school leaving exams raise student achievement and improve how grades are understood in the labor market. *IZA World of Labor*. January 2018: 419.
- Wößmann, L. (2000) *Schooling Resources, Educational Institutions, and Student Performance: The International Evidence*. Working paper no. 983. Kiel, Germany: Kiel Institute of World Economics.
- Wößmann, L. (2003). Schooling Resources, Educational Institutions and Student Performance: The International Evidence. *Oxford Bulletin of Economics and Statistics*, 65(2), 117-170.
- Wößmann, L. (2016) The importance of school systems: Evidence from international differences in student achievement. *Journal of economic perspectives*. 30. 3-32.
- Yeager, D. S., & Walton, G. M. (2011). Social psychological Interventions in Education: They're not magic. *Review of Educational Research*, 81(2), 267–301.

DEEL 2. BIJLAGEN BIJ HET RAPPORT NAAR HEFBOMEN VOOR ONDERWIJSVERBETERING

Door Jaap Scheerens, Bernard Veldkamp en Martina Meelissen

ANNEX 1 Results from PISA

PISA 2009 (reading literacy performance)

Volume IV of the report on PISA 2009, (OECD, 2010) with reading literacy as the core performance domain is titled; “What makes a school successful; resources, policies and practices”. The report provides information on the association of a number of system and school level variables with student achievement, more in particular results concerning *selectivity and stratification (concepts of horizontal and vertical differentiation), decentralization and school autonomy, choice, accountability, educational resources (i.e. teacher salaries) and school climate*.

The concept of vertical differentiation was coined to refer to the differentiation of 15-year old students across levels (grades) in educational systems, and to grade repetition. Horizontal differentiation refers to the degree of stratification of school structures, in terms of different school categories and age of first selection. The percentage of selective schools in the system was added as a third indicator of horizontal differentiation. The results indicate that strong vertical differentiation in terms of grade repetition is negatively associated with performance. In 24 OECD countries and 27 partner countries schools with more repeaters tended to achieve lower scores, after adjustment for student background. Grade repetition was also associated with low equity in the sense of high SES determinacy of performance.

Horizontal differentiation appeared to be unrelated to performance but tended to have a negative association with equity. School level indices of horizontal differentiation and transfer of students to other schools because of low achievement and ability grouping appeared to be negatively associated with performance across countries. Country by country analyses showed a mixed pattern of positive and negative associations of these variables with performance. School systems with low levels of student transfer tended to have more school autonomy in the domain of curriculum and assessment.

School autonomy, free school choice and accountability arrangements are key features of school governance, addressed in volume IV. School autonomy in the domain of curriculum and assessment appeared to be positively associated with performance, while this positive association was not found with respect to autonomy in the use of school resources. Yet, in countries that had specific features of accountability in place, notably the posting of achievement data publicly by schools, school autonomy in the domain of resources did show a positive association with performance.

Across countries, school choice did not show a relationship with performance. Within countries, schools indicating a stronger level of competition tended to do better, but this association appeared to be highly dependent on the socio economic level of the school intake. This might be interpreted as a sign that high competition stimulates social segregation in schooling. Likewise, the existence of private schools appeared to be unrelated to performance across countries, while within countries positive association depended mostly on the socio economic status of the students.

Standard based examinations appeared to be positively associated with performance across OECD countries. On average countries with standard based examinations scored 16 points higher on reading performance than countries that do not have these. No association was found for the use of standardized tests, across countries; within country analyses pointed out that several countries did show a positive association. Use of assessment was positively associated with equity in the sense of relatively low SES determinacy of performance. Country by country analyses provided a somewhat mixed pattern of positive and negative associations as far as assessment is concerned.

Across countries, a positive association between educational resources, i.e. level of teacher salaries, and performance was found. Within school systems this association appeared to be strongly dependent on the socio economic background of the students. Provision of pre-school education had a positive

association with equity. Resource related variables like instruction time, teacher/pupil ratio and expenditure did not show a positive association across OECD countries, but it did for all countries in PISA. The explanation that is suggested is that in industrialized countries resources do not dramatically vary between countries, and within countries, resources do not differ strongly between schools.

PISA 2012 (mathematics)

In the PISA studies, variables that have an interpretation as system level conditions are mostly based on country aggregated data from responses collected at school level, from school principals. Summary results in this section are based on the PISA 2012 report, OECD (2013, Vol. IV). Associations of the variables in question with student performance can be computed at the country level and at school level. The country level correlations, shown in the second column of table 1, although adjusted for an overall measure of wealth (GDP per capita), are to be considered as rather crude measures. When the relationship of these variables are computed for each country separately, within schools, by means of multi-level analyses, more elaborate adjustment for student background characteristics and other school characteristics are possible. The figures shown in the third column reflect the OECD mean of the adjusted school level effects of these variables, expressed as the performance difference associated with a unit change in the independent variables. The overview in Table 1 indicates that sizeable associations found at the between country level are not reproduced when coefficients are based on within country measures, after adjustment for other variables. The only variable where there is a consistent country and school level association is “teacher shortages”. It should be noted that the relatively high correlations for the two indicators of horizontal differentiation (number of school types, and age of first selection for secondary education) are computed with the SES impact on mathematics performance as the dependent variable. These coefficients are interpreted as measures of equity rather than quality. Horizontal differentiation had no significant association with mathematics performance. The information collected in Table 1 also shows that these variables computed at school level, within countries, are only significantly associated with performance in small minorities of countries, while the number of positive associations is not much higher than the number of negative associations. Results in table 1, show a sizeable country level correlation for school autonomy in the curricular domain. It should be noted that this variable is based on principals’ comments on the freedom that schools have in choosing textbooks and assessment methods. This interpretation is closer to autonomy in the domain of classroom management and instruction than to an interpretation in terms of centralized or decentralized curricula. Quality of resources shows a sizeable between country correlation with achievement and teacher shortage a sizeable negative correlation, both associations having the expected signs.

Table 1: Relationships between variables that have a system level interpretation, at country and school level, source PISA 2012, vol IV, OECD, 2014 and PISA 2015, OECD 2016)

	System level correlation with math performance or ESCS impact on performance*	Performance difference per unit change in independent variable, OECD average across countries	Number of OECD countries with a significant positive association	Number of OECD countries with a significant negative association	Number of OECD countries with non-significant association
School autonomy		(-0.01)		(6)	(29)
curriculum	.58	.05	1	1	32
resources	.00	-4.2	3	2	29
Choice					
competition	-.02	-.01	2	1	31
private	.14	1.3	5	1	28
Accountability					
data public	-.21	2.6 (4)	1 (8)	1	32 (27)
ext authority	-.34	-2.1		1	33
written feedback	.24	-.09	2	2	32
Horizontal differentiation					
number of school types	.65*				
age of first selection sec.	.63*				
Resources					
quality educational res.	.58	-.04 (1)	(2)	1 (1)	34 32
teacher salaries	.31				
Teacher training					
university qualification	-.15	.09 (0.1)	4 (4)	2 (1)	28 (30)
teacher shortage	-.41	-1.8 (-2)	(1)	3 (4)	31 (30)
professional development	.01	0.0	1		33
OTL formal mathematics		48.9	33		0 (1 missing)

Note: Figures in the second column are country level correlations. Figures in the third column are performance differences per unit change in the independent variable, obtained from multi-level analyses, controlling for student background and other school characteristics. Figures in bold indicate statistical significance at the .05 level. OTL is "opportunity to learn". Sources, OECD, 2014, Table IV.1.3 (system level correlations) and Table IV 1.12c (within country associations). Entries in brackets are results of PISA 2015

The country level correlations for accountability and teacher training are insignificant and sometimes, counter to expectations, show negative signs. The coefficient on opportunity to learn in academic mathematics, suggest that content covered makes an important difference, within countries (and probably also between countries). However, some caution should be exercised regarding this unusually high effect-size for a (*any*) teaching level variable in international studies. For one thing the OTL coefficient was computed for a limited set of items about formal mathematics. It could also matter that the PISA OTL measure depended on responses from students, whereas OTL measures based on TIMSS, and which usually show much more modest effect sizes, are based on teacher judgments (Luyten & Scheerens, 2021, in press).

PISA 2015 (*science*)

For those variables that were analyzed in a comparable way in PISA 2015, results in Fig. 5.x are listed in brackets. The measure for school autonomy is based on an overall index of autonomy. The impact of accountability based on published school data was slightly stronger, compared to PISA 2012. Results on quality of educational resources, teacher qualifications and teacher shortages were quite similar between these two waves of PISA.

The overview of results from PISA 2015 for the system level variables addressed in this chapter is as follows:

Autonomy

After accounting for the socio-economic profile of students and schools, there is no association between school autonomy and student performance in science, on average across OECD countries (p.37)

Functional decentralization

“PISA 2015 offers a nuanced picture of the relationship between greater school autonomy and students’ performance, which seems to depend not only on the particular areas of school management delegated to principals and teachers, but also on how these areas are related to certain accountability measures and to the capacity of local actors. In particular, students score higher in science when principals exercise greater autonomy over resources, curriculum and other school policies, but especially so in countries where achievement is tracked over time or posted publicly or when principals show higher levels of educational leadership”. (p. 45)

Private public

On average, across OECD countries and in 32 education systems, student enrolled in public schools score lower in science than students in private school so. But, as has been noted in previous PISA reports, this is no longer the case after accounting for socio-economic status (p. 39).

Accountability and evaluation

On average across OECD countries, students enrolled in schools that post achievement publicly, score 4 points higher, than schools not following this practice (p. 353, table II.4.30)

Stratification

“Selecting students into different programs or schools, especially when students are young, is strongly associated with less academic achievement across schools and less equity in science performance” (p 40).

However, as referred to previously, the Netherlands does not conform to this overall picture, which OECD attributes to other aspects of the educational context (OECD, 2015, p 173-175).

Other interesting results from PISA 2015 are that no significant effect is found for the level of teacher salaries, relative to per capita national income, that educational leadership has a slightly negative effect and that teacher cooperation in professional exchange shows an effect of 9 score points (p. 43).

PISA 2018 (reading)

PISA 2018, with reading literacy as the main subject, does not address school autonomy and functional decentralization. The nuanced effects of these structural governance variables appeared to be quite consistent across previous PISA editions. Results on private/public, accountability and evaluation and stratification restate conclusions reached in previous editions; PISA 2015 included. Additional areas highlighted in the report of PISA 2018, particularly vol V titled “Successful policies, Successful schools” are: the impact of resources, school choice, educational technology and professional development. In the following overview summary conclusions in all major areas are cited from the report.

Major findings on the effect of system level levers, cited from PISA 2018, Vol V.

Budget

Amongst the countries and economies whose cumulative expenditure per student was under USD 50 000 (the level of spending in 24 countries/economies), higher expenditure on education was significantly associated with higher scores in the PISA reading test. But this was not the case amongst countries and economies whose cumulative expenditure was greater than USD 50 000, which include most OECD countries (Figure II.6.2). It seems that for this latter group of countries and economies, factors other than the level of investment in education are better predictors of student performance 97

Salaries

Higher salaries can help school systems attract the best candidates to the teaching profession, and signal that teachers are regarded and treated as professionals. But paying teachers well is only part of the equation. The relationship between reading performance and teachers’ salaries relative to per capita national income was statistically significant across OECD countries and across PISA-participating countries and economies in 2018 (Figure V.4.8). However, the correlation was entirely driven by Mexico and Lebanon, two countries where teachers’ PPP-corrected per capita salaries are higher than those in other countries/economies. After excluding these outliers, the relationship was not statistically significant. This finding suggests that other factors, such as the quality of teaching, may be more closely associated with students’ performance at the system level. 97

Human resources

The percentage of teachers fully certified by an appropriate authority was positively correlated with mean performance in reading, even after accounting for per capita GDP, across OECD countries. Differences in teacher certification accounted for about 16% of the differences in mean reading performance across all countries and economies. Moreover, the percentage of full-time teachers was associated with greater equity in reading performance across all countries and economies, even after accounting for per capita GDP. 196

Technology

After accounting for students’ and schools’ socio-economic profile, in 11 countries/economies, students in schools where a larger share of the computers available to students for educational purposes is connected to the Internet scored higher in reading; but in 7 countries/economies students in such schools scored lower 113

While results show that the proportion of portable computers in school is largely unrelated to students' reading scores, having high-speed Internet access at school is associated with better performance, even after accounting for socio-economic status. 195

Stratification

In PISA 2018, education systems with a larger number of education programs available to 15-year-olds generally showed lower mean performance in reading and less equity in reading performance 71

Offering fewer instructional tracks was correlated not only with student performance but also, and more strongly, with greater equity in performance, as shown in Figure V.3.7. After accounting for per capita GDP, a correlation was observed between the number of instructional tracks and equity in performance in reading, mathematics and science (Table V.B1.3.12) 81

Students' age at first selection into different programs was not consistently correlated to mean reading performance. However, selecting students into different programs at an earlier age was correlated with less equity in reading performance, even after accounting for per capita GDP, across OECD countries, and across all countries/economies (Figure V.3.9). As shown in Figure V.3.9, differences in the age at first selection accounted for 46% of the differences in equity in reading performance across OECD countries, and for almost 20% of the differences in equity in reading performance across all countries and economies. 84

Private Public

After accounting for students' and schools' socio-economic profile, students in public schools scored higher in reading than students in private schools, on average across OECD countries (by 14 score points, in favour of public schools) and in 19 education systems (ranging from 13 score points higher in Indonesia to 117 points higher in Serbia). At the system level, across all countries and economies, school systems with larger shares of students in private-independent schools tended to show lower mean performance in reading, mathematics and science, after accounting for per capita GDP. 158

Choice

No clear system-level patterns were observed in the relationships between various indicators of school competition, on the one hand, and performance and equity in reading performance, on the other. 185

Professional development

The relationship between teachers' participation in professional development activities and students' performance in reading is weak in most PISA-participating countries and economies (Table V.B1.4.8). After accounting for the socio-economic profile of students and schools, in nine education systems, students scored higher in reading when more teachers in their school had participated in professional development activities; in ten other systems, students scored lower in reading when their teachers had participated in such activities. 95

Curriculum alignment

High-performing countries and economies tended to have more teacher mentoring on the school's initiative. In those systems, more schools implemented a standardized policy for reading-related subjects taught at school (including a school curriculum with shared instructional materials, and staff development and training) based on district or national policies. 170

Fig 5.1 Selected results PISA 2018

5.4.3.1 Samenvatting van de effecten van systeemvariabelen in PISA (2009, 2012, 2015 en 2018)

De PISA-resultaten worden gerapporteerd voor de kernvariabelen die ook gebruikt zijn om de resultaten van effectiviteitsonderzoek samen te vatten in paragraaf 5.4.2. Hierbij ga ik achtereenvolgens in op de volgende rubrieken: Privéscholen en keuzevrijheid, schoolautonomie en functionele decentralisatie, evaluatie en accountability en stratificatie van scholenstelsels. Vervolgens behandelen als aanvullende rubrieken: Financiële randvoorwaarden, het scholingsniveau van leerkrachten, het effect van deelname aan professionele ontwikkelingsactiviteiten, technologische voorzieningen en “curriculum alignment”.

Privéonderwijs en keuzevrijheid

De Pisa-resultaten wijzen over het algemeen uit dat verschillen in scores tussen privéscholen en openbare scholen verdwijnen (niet meer significant zijn) nadat gecontroleerd is voor de samenstelling van de schoolbevolking qua sociaaleconomische status (PISA 2012, 2015 en 2018). Competitie tussen scholen heeft geen effect of een effect dat verdwijnt na controle voor sociaaleconomische status (PISA 2009, 2012, 2018).

Schoolautonomie en functionele decentralisatie

In PISA is enkele keren aangetoond dat autonomie in het curriculaire domein een gunstig effect heeft (PISA 2009 en PISA 2012). Hierbij moet worden opgemerkt dat de invulling van “curriculair” hier bepaald is als de vrijheid van docenten om leermethoden en assessment vormen in de klas te bepalen en dus niet in de betekenis van het bestaan van een voorschrijvend nationaal curriculum dan wel een meer “open” curriculum. Van autonomie op het gebied van het zelf managen van financiële en materiële middelen werd vastgesteld dat die een positief effect heeft, maar alleen als dit samengaat met evaluatieve en accountability voorzieningen (gegevens uit PISA 2009 en 2012). In PISA 2015 werd op basis van een meer algemene autonomie index vastgesteld dat deze variabele geen effect had op prestaties.

Evaluatie en accountability

In PISA 2009 werd vastgesteld dat landen die een op standaarden gericht centraal examen hadden gemiddeld 16 punten hoger scoorden bij lezen. De uitkomsten van PISA 2012 lieten een negatief effect zien van het ter beschikking stellen van evaluatiegegevens door scholen aan een “externe autoriteit”. PISA 2015 rapporteert dat het openbaar maken en publiceren van evaluatiegegevens van scholen een positief effect heeft (gemiddeld een score van 4 punten hoger). In de rapportage over PISA 2018 blijkt dit effect 5 punten te zijn. In Pisa 2018 werd verder een positief effect van assessment en feedback vastgesteld op gelijkheid (equity).

Stratificatie

Voor stratificatie worden twee indicatoren gebruikt: het aantal schoolcategorieën in het voortgezet onderwijs en de leeftijd waarop voor een bepaalde vorm van voortgezet onderwijs moet worden gekozen. De algemene trend is dat de prestaties lager zijn naarmate er meer school categorieën zijn en er op jongere leeftijd gekozen moet worden. Nederland scoort op beide indicatoren relatief hoog. De positieve effecten van meer geïntegreerde scholenstelsels en latere selectie gaan zowel op wat betreft hogere prestatieniveaus als voor ongelijkheid, maar het meest consistent voor ongelijkheid (resultaten uit PISA 2009, PISA, 2012, PISA 2015 en PISA 2018)

Budget en salarissen

Alleen voor landen met een investering van minder dan 50.000 USD maakt de hoogte van het onderwijsbudget uit voor de prestaties. Bij rijkere landen, cq de OECD-landen is er geen effect (PISA 2018). Uit PISA 2009-gegevens bleek dat er tussen landen een positief effect was voor financiële middelen, salaris inbegrepen, maar dat dit effect binnen landen verdween na controle voor sociaaleconomische status. Salaris bleek in PISA 2018 geen effect te hebben, nadat extreme uitkomsten van 2 landen verwijderd waren.

Het scholingsniveau van leerkrachten

Uit de resultaten van PISA 2012 bleek dat het percentage docenten met een universitaire opleiding geen significant effect op de leerprestaties had. Het percentage bevoegde leraren bleek in PISA 2018 een positief effect te hebben op de prestaties. Verder bleek dat het percentage *full-time* leerkrachten in een land een positief effect te hebben op gelijkheid (*equity*).

Deelname aan professionele ontwikkelingsactiviteiten

Uit de resultaten van PISA 2012 en 2018 bleek dat van deze deelname geen effect op de prestaties kon worden vastgesteld.

Lerarentekorten

Uit de resultaten van PISA 2012 en PISA 2018 bleek dat door scholen ervaren tekorten een negatief effect op de prestaties hadden.

Curriculum alignment

Curriculum alignment is als zodanig niet systematisch onderzocht in PISA (Ik kwam de term voor het eerste tegen in het rapport over PISA 2018 (Vol V). De uitkomsten uit PISA 2012 over OTL (gelegenheid om te leren) waren weliswaar spectaculair, maar zijn eenmalig en discutabel. In het rapport van PISA 2018 is er sprake van “a standardized policy for reading related subjects taught at school”; maar de effectiviteit ervan is niet onderzocht.

ANNEX 2 Results from TIMSS and other international studies

TIMSS 2019 (Mathematics)

The TIMSS 2019 reports provide information on the association of some system and school level variables with student mathematics achievement. An overview of the relevant international tables of TIMSS 2019 is presented in Table 1. Some factors are not measured/reported since TIMSS-2015, therefore the results of 2011 were used for these factors.

Table 1 Overview of factors in this study measured in TIMSS

Factor	In TIMSS measured as:	Source	Remarks
Opportunity to learn	Students Taught the TIMSS Mathematics Topics	https://timss2019.org/reports/classroom-contexts/	Not reported with country average achievement scores
	Instructional Time in Mathematics	https://timss2019.org/reports/classroom-contexts/	Not reported with country average achievement scores
Teachers' pedagogical content knowledge	TIMSS 2011: Teachers Feel "Very Well" Prepared to Teach TIMSS Mathematics Topics, Fourth Grade	https://timssandpirls.bc.edu/timss2011/downloads/T11_IR_Mathematics_FullBook.pdf	Not reported with country average achievement scores
Support	Instructional Clarity in Mathematics Lessons – Students' Reports	https://timss2019.org/reports/classroom-contexts/	Reported with country average achievement scores
Safe and orderly climate	Disorderly Behavior During Mathematics Lessons – Students' Reports	https://timss2019.org/reports/classroom-contexts/	Reported with country average achievement scores
	Student Bullying – Students' Reports	https://timss2019.org/reports/classroom-contexts/	Reported with country average achievement scores
	Safe and Orderly School – Teachers' Reports	https://timss2019.org/reports/home-contexts/	Reported with country average achievement scores
	School Discipline – Principals' Reports	https://timss2019.org/reports/home-contexts/	Reported with country average achievement scores
Parental involvement	Home Early Literacy and Numeracy Activities Before Primary School – Parents' Reports	https://timss2019.org/reports/home-contexts/	Reported with country average achievement scores
Cooperation between teachers	TIMSS 2011 Collaborate to Improve Teaching, Fourth Grade	https://timssandpirls.bc.edu/timss2011/downloads/T11_IR_Mathematics_FullBook.pdf	Reported with country average achievement scores
Achievement orientation	School Emphasis on Academic Success – Principals' Reports	https://timss2019.org/reports/home-contexts/	Reported with country average achievement scores

The international tables only provide descriptives (Mullis et al, 2012; Mullis et al, 2020) and do not report results of correlational analyses or regressions analyses. TIMSS reports these descriptives by using indexes with categories (for example 'low'- 'moderate'- 'high'). In most cases, countries' average mathematics achievement is reported for each index-category separately. However, this information only suggests that these factors might or might not be related to students' achievement. Without further analyses on the actual data of each country, it cannot be concluded that these factors are related or not in the different countries.

The overview of results from TIMSS for the relevant system and school level variables addressed in this study is as follows:

Support

The effect of support on mathematics performance is examined by measuring instructional clarity and relating the score to the average scale score at country level. Students are classified as high, medium or low clarity of instruction. For each participating country, the percentage of students in each category is reported together with the average scale score for these students. No further analysis of the results is given. Within each country, it can be seen that higher instructional clarity at country level is associated with higher average performance. A more detailed analysis of the results would have to reveal whether this relation can be found at the level of individual students.

Safe and orderly climate

Safe and orderly climate is measured in TIMSS 2019 with four scales: dealing with disorderly behavior during the lessons, bullying, safe-and-orderly school and school discipline. Especially the scales related to disorderly behavior and bullying show on averaged quite a difference in mathematics achievement when students that classify as high on these scales are compared to the students that classify in the low category. Since only descriptives have been reported, secondary analyses have to reveal whether a relation can be found between climate and achievement.

Parental involvement

The involvement of parents is measured with a scale about early literacy and numeracy activities before primary school. Students were scored according to their parents reports. They scored as (almost) never, sometimes, or often participating in these activities. On average, the students who participated often in these kind of activities before primary school showed a higher mathematics achievement than those who participated sometimes. Only a small percentage of students participated (almost) never in these kinds of activities. To demonstrate the effectiveness of participating in these activities before primary school, further analyses have to be conducted.

Cooperation between teachers

This scale was last measured in TIMSS 2011. Based on the reports of the teachers, students are classified as having 'very collaborative', 'collaborative', or 'somewhat collaborative' teachers. On average, cooperation between teachers had hardly any effect on average math achievement within a country. From the descriptives, it can be seen that in some countries, students who attended schools where teachers somewhat collaborated demonstrated a lower mathematics achievement than students who attended school where teachers were very collaborative. But other countries showed a reversed effect.

Achievement orientation

Based on reports of principles, students are classified as attending schools with a 'very high emphasis', 'high emphasis', or 'medium emphasis' on academic success. For each country, the percentage of students in each category, as well as their average mathematics achievement is presented. For quite a few countries, it can be seen emphasis on academic success is somewhat related to mathematics achievement in TIMSS. When different countries are compared, it is not clear whether countries that emphasize academic success score higher than countries that do not emphasize academic success. Secondary analysis is needed to answer these kinds of questions.

Additional literature search

The recent literature (since 2015) was scanned for meta-analyses related to PISA and TIMSS and the impact of various kinds of policy measures on student performance. We searched for literature on the following keywords:

- Opportunity to learn
- Curriculum alignment
- Teachers' pedagogical content knowledge
- Direct teaching
- Cognitive activation
- Classroom management
- Support
- Safe and orderly climate
- Parental involvement
- Educational leadership
- Cooperation between teachers
- Achievement orientation
- Assessment for learning
- High stakes testing
- Examinations
- Accountability
- School autonomy
- Adaptive teaching
- School as learning organization
- Professional learning communities
- feedback

In this limited literature review, we did not get too many hits. For an overview of the 29 studies found, see the reference list. The results can be summarized in the following way.

Impact of various policy measures (based on recent papers)

Overall findings

Based on thirty years of international large scale assessments, one has to be careful in trying to predict the impact of various policy measures on the performance of individual students. Hopfenbeck et al (2018) emphasize the need for caution when using this kind of research to inform educational policy, because no clear claims seem to be supported consistently. Deng and Gopinathan (2016) mention that high standards

in teacher recruitment, effective programmes of teacher preparation, opportunities for professional development and school leadership are certainly important and contribute to the 'quality' of teachers, but they cannot determine what teachers do in classrooms. While system characteristics – such as high academic standards, expectations, accountability mechanisms, and so forth – could promote the academic performance of students, they are embedded in and work hand in hand with other cultural and institutional aspects of an education system – including the curriculum system, the examination system, the instructional system, conventional beliefs and expectations on schooling, among others.

Opportunity to learn

Scheerens (2016) studied the association between measures of OTL and instructional alignment with student performance. He found that the OTL significantly affected the performance, but that effect size was relatively small ($d = 0.30$). In order to finetune the results of Scheerens (2016), Lamain (2018) studied a narrower definition of OTL, as well as, the influence of study characteristics. But this gave no explicit explanation for the modest proportion of significantly positive OTL. Based on data from TIMSS 1999 – 2007, Lin (2018) found a positive effect of the instructional time component with mathematics achievement. This is in line with Scheerens (2016), Avieko (2018) concluded, bases on TIMSS 2015 data, that the findings indicate that the Opportunities to Learn algebra are context specific and align with the expectations of the curriculum guides from the governing bodies within the countries. That study also demonstrated that students who learned particular algebra topics in the early grades had significantly higher algebra scores in the TIMSS 2015 assessment.

pedagogical content knowledge

Ulferts (2019) showed the relevance of general pedagogical knowledge for successful teaching: She reported significant effects for teaching quality and student outcomes (Cohen's $d = .64$ and $.26$), indicating that more knowledgeable teachers achieve a three-month additional progress for students. Therefore she concluded that general pedagogical knowledge is a crucial resource for teaching.

Direct teaching

Holzberger et al (2020) showed that instructional practices are significantly related to performance ($r = .13$). Direct teaching turned out to be most effective according to several studies. Cordero and Gil-Izquierdo, M. (2018) found that traditional teaching methods have a positive influence on students' proficiency in mathematics, while the implementation of more innovative active learning strategies seem to have a negative impact on student achievement. Forbes et al (2020) revealed fundamental differences in the frequency and nature of inquiry-based instruction reported by students associated with the highest and lowest levels of science achievement. High levels of inquiry-based teaching are relatively uncommon and associated with lower student science achievement while higher levels of student science achievement are associated with relatively more teacher-directed forms of inquiry.

Support

We only found one study related to support. The results indicate that perceived teacher support is beneficial to student learning by fostering academic self-concept and academic enjoyment. Ma, Luo & Xiao (2021).

Safe and orderly climate

A small positive effect was found for a safe and orderly climate at school and in the classroom. Lee and Stankov (2018) studied both TIMSS and PISA. For TIMSS, they demonstrated that student feeling of school safety was positively associated with mathematics performance. Based on PISA data, they also found a positive relation a disciplinary climate with student performance. Wang, Gao, & Shang, (2018) found a non-significant positive relation between culture and performance in PISA 2006 data. Holzberger showed that the classroom climate had a small positive relation with performance. Finally You, Park and Delgado (2020) showed that climate is a significant predictors of all students' performance in US.

Parental involvement

For parental involvement, some mixed results were obtained. In East Asian countries a positive association between parental involvement and achievement was found by Kim (2020). The strength of the relation was highest for academic socialization, followed by home involvement and school involvement, similar to previous meta-analyses. Reparaz and Sotés-Elizalde (2019) studied parental involvement more into detail. They found that spending time just talking, being interested, discussing child's behavior initiated by the teacher, attending meetings and talking about support with the teacher were positive ways of parental involvement. Helping with homework and discussing child's behavior with the teacher initiated by the parents are negatively related to performance. Participating in committees and parent council did not affect the performance. These findings are confirmed by Hartas (2015), who showed that less instrumental, more subtle forms of parental involvement such as parent-child conversations about topical social issues emerged as the strongest predictor for continuing parental literacy support at age 15.

Educational leadership

Educational leadership can be studied from various perspectives. Overall, Karadag (2020) reported that educational leadership has a medium-level effect on students' achievement. If you take into account the difference in culture, educational leadership has comprehensive effect on student achievement in vertical-collectivist cultures (e.g., in Asian) than horizontal-individualistic cultures (e.g., in USA). As was expected, the most comprehensive effect among leadership theories was found in instructional leadership. Wu, Gao & Shen (2020) found (a) that principals' rating of their own overall leadership was statistically negatively related to student achievement, (b) that among the four subscales of principal self-rated leadership, instructional leadership was statistically positively related to student achievement, while leadership for teacher development was negatively related to student achievement and (c) that principal leadership's effect on student achievement was moderated by school contextual variables. School size had a small negative effect. It could be mentioned that instructional leadership was also in this study found to be positively related to performance. Besides studying principal leadership from the principal's perspective, Wu et al (2020) studied principal leadership from the teacher's perspective and they found that teachers' rating of principal leadership had a direct and positive relationship with student science achievement. They also found that teachers' job satisfaction and collaboration were positively associated with principal leadership but did not mediate the relationship between principal leadership and student science achievement. Finally, they noted that principal leadership's direct effect on science achievement was moderated by school size.

School autonomy

With respect to the impact of school autonomy, we also found one source. Suggett (2015) mentions that the data points to autonomy as a key and necessary component of a mature and high-performing system, as it is in other areas of public administration. But Suggett (20105) also emphasized that accountability

and leadership are important. She ends with an interesting observation that more rather than less systemic support is needed for the potential of school autonomy to be realized.

Adaptive teaching

As was already mentioned with the overall findings, it matters what teachers do in the classroom. Conn (2017) reports that that programs that alter teacher pedagogy or classroom instructional techniques had an effect size approximately 0.30 standard deviations greater than all other types of programs combined. Limited evidence further suggests that pedagogical programs that employed adaptive instruction or teacher coaching were particularly effective. Oliver et al (2019) found that students who reported experiencing high frequencies of inquiry strategies in their classrooms consistently evidenced lower levels of scientific literacy. They also found a strongly positive association between the frequency of teacher-directed and adaptive teaching strategies and students' scientific literacy. Adaptive teaching is often accompanied by the use of digital tools. Hillmayr et al (2020) reported, based on, amongst other studies, PISA 2015 and 2018, that digital tool use had a positive effect on student learning outcomes ($g = 0.65, p < .001$). The provision of teacher trainings on digital tool use significantly moderated the overall effect. Use of intelligent tutoring systems or simulations such as dynamic mathematical tools was significantly more beneficial than hypermedia systems. On a descriptive level, the effect size was larger when digital tools were used in addition to other instruction methods and not as a substitute.

Feedback

Only one study was found on a meta-analysis that included PISA/TIMMS and feedback. This study reported that Elaborated feedback was particularly more effective than Knowledge of response and Knowledge of correct response for higher order learning outcomes. Larger effect sizes were found for mathematics compared with social sciences, science, and languages. Effect sizes were negatively affected by delayed feedback timing (Van der Kleij et al, 2015).

Reference list

- Ayieko, R. A. (2018). The Relationship between Opportunities to Learn Algebra and Students' Algebra Achievement: A Comparative Study. *Global Education Review*, 5(3), 29-56.
- Carnoy, M., Khavenson, T., & Ivanova, A. (2015). Using TIMSS and PISA results to inform educational policy: a study of Russia and its neighbours. *Compare: A Journal of Comparative and International Education*, 45(2), 248-271.
- Conn, K. M. (2017). Identifying effective education interventions in sub-Saharan Africa: A meta-analysis of impact evaluations. *Review of Educational Research*, 87(5), 863-898.
- Cordero, J. M., & Gil-Izquierdo, M. (2018). The effect of teaching strategies on student achievement: An analysis using TALIS-PISA-link. *Journal of Policy Modeling*, 40(6), 1313-1331.
- Deng, Z., & Gopinathan, S. (2016). PISA and high-performing education systems: Explaining Singapore's education success. *Comparative Education*, 52(4), 449-472.
- Fan, H., Xu, J., Cai, Z., He, J., & Fan, X. (2017). Homework and students' achievement in math and science: A 30-year meta-analysis, 1986–2015. *Educational Research Review*, 20, 35-54.
- Forbes, C. T., Neumann, K., & Schiepe-Tiska, A. (2020). Patterns of inquiry-based science instruction and student science achievement in PISA 2015. *International Journal of Science Education*, 42(5), 783-806.
- Lee, J., & Stankov, L. (2018). Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences*, 65, 50-64.
- Håkansson, J. (2015). Structured teaching and classroom management—The solution for the decline of Swedish school results? Conclusions drawn from a comparative meta-synthesis of teaching and learning. *Teachers and Teaching*, 21(5), 584-602.
- Hartas, D. (2015). Patterns of Parental Involvement in Selected OECD Countries: Cross-National Analyses of PISA. *European Journal of Educational Research*, 4(4), 185-195.
- Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education*, 153, 103897.
- Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., & Baird, J. A. (2018). Lessons learned from PISA: A systematic review of peer-reviewed articles on the programme for international student assessment. *Scandinavian Journal of Educational Research*, 62(3), 333-353.
- Holzberger, D., Reinhold, S., Lüdtke, O., & Seidel, T. (2020). A meta-analysis on the relationship between school characteristics and student outcomes in science and maths—evidence from large-scale studies. *Studies in Science Education*, 56(1), 1-34.

- Karadag, E. (2020). The effect of educational leadership on students' achievement: A cross-cultural meta-analysis research on studies between 2008 and 2018. *Asia Pacific Education Review*, 21(1), 49-64.
- Kim, S. W. (2020). Meta-analysis of parental involvement and achievement in East Asian countries. *Education and Urban Society*, 52(2), 312-337.
- Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of educational research*, 85(4), 475-511.
- Lamain, M. (2018). *The effects of opportunity to learn on students' achievement: a meta-analysis* (Master's thesis, University of Twente).
- Ma, L., Luo, H., & Xiao, L. (2021). Perceived teacher support, self-concept, enjoyment and achievement in reading: A multilevel mediation model based on PISA 2018. *Learning and Individual Differences*, 85, 101947.
- Oliver, M., McConney, A., & Woods-McConney, A. (2019). The efficacy of Inquiry-Based instruction in science: A comparative analysis of six countries using PISA 2015. *Research in Science Education*, 1-22.
- Reparaz, C., & Sotés-Elizalde, M. A. (2019). Parental involvement in schools in Spain and Germany: Evidence from PISA 2015. *International Journal of Educational Research*, 93, 33-52.
- Rios, J. (2021). Improving test-taking effort in low-stakes group-based educational testing: A meta-analysis of interventions. *Applied Measurement in Education*, 1-22.
- Scheerens, J. (Ed.). (2016). *Opportunity to learn, curriculum alignment and test preparation: A research review*. Springer.
- Scheerens, J., Lamain, M., Luyten, H & Noort, P. (2017). Opportunity to Learn, Curriculum Alignment and Test Preparation. A Research Review. Houten: Springer.
- Suggett, D. (2015). School autonomy: Necessary but not sufficient. *Evidence base: a journal of evidence reviews in key policy areas*, 2015(1), 1.
- Ulferts, H. (2019). The relevance of general pedagogical knowledge for successful teaching: Systematic review and meta-analysis of the international evidence from primary to tertiary education.
- Wang, Z., Gao, S., & Shang, Y. Y. S. G. Q. (2018). Influencing Factors of PISA Science Performance: a qualitative meta-analysis. *Journal of Science Education*, 19(2), 107-121.
- Wu, H., Gao, X., & Shen, J. (2020). Principal leadership effects on student achievement: a multilevel analysis using Programme for International Student Assessment 2015 data. *Educational Studies*, 46(3), 316-336.

Wu, H., Shen, J., Zhang, Y., & Zheng, Y. (2020). Examining the effect of principal leadership on student science achievement. *International Journal of Science Education, 42*(6), 1017-1039.

Yorke, J., & Vidovich, L. (2016). A Meta-analysis of Policy Processes. In *Learning Standards and the Assessment of Quality in Higher Education: Contested Policy Trajectories* (pp. 217-257). Springer, Cham.

You, H. S., Park, S., & Delgado, C. (2020). A closer look at US schools: What characteristics are associated with scientific literacy? A multivariate multilevel analysis using PISA 2015. *Science Education*.

ANNEX 3 De rol van technologie en AI in het onderwijsproces

Een van de factoren waar de laatste jaren sterk op ingezet is, is het gebruik van technologie in het onderwijsproces. Daarbij is in het bijzonder aandacht besteed aan de meerwaarde die artificial intelligence (AI) kan hebben. De vraag is in hoeverre technologie succesvol ingezet kan worden als ondersteuning van de docent.

In tabel 3.2 worden onder *Interactive strategies* een aantal factoren genoemd waarbij techniek een grote rol kan spelen. In het bijzonder geldt dit voor de factoren:

- klassenmanagement;
- adaptief onderwijs;
- activerend onderwijs;
- uitdagende leeromgeving.

Technologie kan op veel verschillende manieren ingezet worden. Het kan gebruikt worden als hulpmiddel in de presentatie, als middel om leerlingen te activeren, voor het geven van feedback of om het leerproces adaptief in te richten. Er zijn veel studies en meta-analyses uitgevoerd om de impact van de inzet van technologie te evalueren. Chauhan (2019) en Garcon & Azevedo (2019), bijvoorbeeld, laten in hun meta-analyses zien dat het gebruik van verschillende vormen van technologie door leraren gemiddeld genomen een middelgroot effect hebben op de leeropbrengst. Er blijken wel grote verschillen te zitten tussen de effectiviteit van verschillende technologieën en voor de verschillende vakgebieden. De inzet bij STEM-vakken heeft een groot effect, terwijl er slechts sprake is van een klein effect bij de sociale vakken. Ook is de duur van inzet en het klimaat van belang bij de effectiviteit.

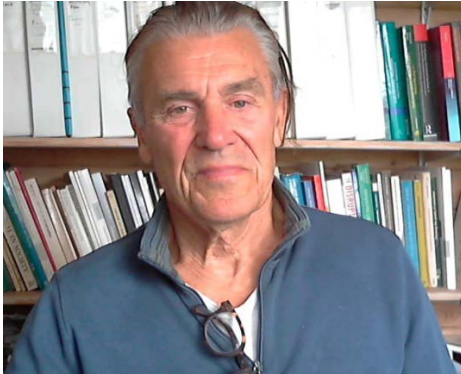
Artificial intelligence gaat over het vermogen van een systeem om externe gegevens correct te interpreteren, om te leren van deze gegevens en om deze lessen te gebruiken om specifieke doelen en taken te verwezenlijken via flexibele aanpassing (Kaplan & Haenlein, 2019). Horst (2019), Marr (2018) geeft aan dat de inzet van AI in het onderwijs zich concentreert op hoe docenten AI gebruiken, hoe AI en learning analytics ingezet kan worden voor adaptief leren en gepersonaliseerd onderwijs, hoe AI-leermateriaal dat vrij beschikbaar is kan ontsluiten voor docenten en leerlingen en hoe AI de administratieve taken van docenten kan verlichten. De eerste meta-analyses (e.g. Ma et al., 2014) laten middelgrote effect sizes zien. Onderzoek heeft uitgewezen dat onderwijskundige toepassingen van AI steeds vaker ontwikkeld worden om geïntegreerd te worden in een didactische aanpak waarbij de docent en de leerling adaptief en dynamisch gebruik maken van de kennis en aanbevelingen van AI-systemen. Holstein, Aleven, & Rummel (2020) geven een theoretisch framework waarbinnen de inzet van AI door docenten vormgegeven kan worden in een hybride aanpak. Wel brengt de inzet van AI allerlei uitdagingen met zich met op het gebied van ethiek, privacy en beveiliging (Veldkamp et al, 2021). Davies et al. (2020) adviseren om het gebruik van AI in het onderwijsproces grondig te doordenken en evalueren voor het op grote schaal ingezet wordt.

Referenties

- Chauhan, S. (2017). A meta-analysis of the impact of technology on learning effectiveness of elementary students. *Computers & Education, 105*, 14-30.
- Davies, H. C., Eynon, R., & Salveson, C. (2020). The Mobilisation of AI in Education: A Bourdieusean Field Analysis. *Sociology*, 0038038520967888.
- Garzón, J., & Acevedo, J. (2019). Meta-analysis of the impact of Augmented Reality on students' learning gains. *Educational Research Review, 27*, 244-260.
- Holstein, K., Alevan, V., & Rummel, N. (2020, July). A Conceptual Framework for Human–AI Hybrid Adaptivity in Education. In *International Conference on Artificial Intelligence in Education* (pp. 240-254). Springer, Cham.
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons, 62*(1), 15-25.
- Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of educational psychology, 106*(4), 901.
- Marr, B. (2018). How Is AI Used In Education--Real World Examples Of Today And A Peek Into The Future. *Forbes, Forbes Magazine, 25*.
- Veldkamp, B., Schildkamp, K., Keijsers, M., Visscher, A., & de Jong, T. (2021). Big Data Analytics in Education: Big Challenges and Big Opportunities. *International Perspectives on School Settings, Education Policy and Digital Strategies: A Transatlantic Discourse in Education Research, 266*.

Cv's project team

Cv Jaap Scheerens



Jaap Scheerens studeerde pedagogiek (onderwijskunde) en psychologie (methodenleer) en promoveerde in 1983 cum laude aan de Universiteit van Groningen. Het grootste deel van zijn beroepsloopbaan was hij hoogleraar onderwijskunde aan de Universiteit Twente. Van 1986 tot 2011 gaf hij leiding aan de vakgroep Onderwijsorganisatie en -management. Het onderzoeksprogramma van de vakgroep, op het terrein van onderwijseffectiviteit, werd in vier achtereenvolgende onderzoeksvisitatie als zeer goed tot excellent beoordeeld. Tussen 1988 en 1998 was hij directeur van het facultaire onderzoeksinstituut OCTO, en van 1994 tot en met 2002 was

hij wetenschappelijk directeur van de onderzoeksschool ICO. Hij heeft gedurende zijn gehele universitaire loopbaan tal van internationale onderzoeksprojecten uitgevoerd, onder meer voor de OECD, de Wereld Bank, UNESCO en de Europese Unie. Hij publiceerde tot nu toe een twintigtal boeken en 100 wetenschappelijke artikelen, voornamelijk op het terrein van onderwijseffectiviteit en verschillende vormen van evaluatie in het onderwijs. Hij is als professor emeritus te gast bij de Universiteit Twente, werkt momenteel op *ad hoc* basis als senior onderzoeker bij Oberon, en is lid van de wetenschappelijke adviesraad van INVALSI, Italië.

10 publicaties Jaap Scheerens

Scheerens, J., Luyten, H., and Van Ravens, J. (2011). *Perspectives on educational quality. Illustrative outcomes on primary and secondary schooling in the Netherlands*. Dordrecht, Heidelberg, New-York, London: Springer

Scheerens, J. (Ed.). (2012) *School leadership effects revisited: Review and meta-analysis of empirical studies*. Dordrecht: Springer.

Scheerens, Jaap (2013) *OECD review on evaluation and assessment frameworks for improving school outcomes: update to the country background report for the Netherlands*.

Scheerens, J. (Ed.). (2014). *Effectiveness of time investments in education: Insights from a review and meta-analysis*. Dordrecht: Springer.

Scheerens, Jaap (2015) Theories on educational effectiveness and ineffectiveness. *School Effectiveness and School Improvement*, 26 (1). 10 - 31. ISSN 0924-3453

Scheerens, J. (2016) *Educational effectiveness and ineffectiveness: A critical review of the knowledgebase*. Dordrecht: Springer

Scheerens, J. (2017a) Ed. *Opportunity to learn, curriculum alignment and test preparation. A research review*. Dordrecht, New York, London: Springer

Jaap Scheerens & Gary N. Marks (2018): Malleability in educational effectiveness: what are realistic expectations about effect sizes? Introduction to the special issue, *Educational Research and Evaluation*, DOI: 10.1080/13803611.2017.1455280

Scheerens, J. (2019) School effectiveness and school organization. In: Connolly, M., Eddy-Spicer, D.H., James, C., and Kruse, S.D. *The Sage handbook of school organization*. London: Sage.

Scheerens, J., Van der Werf, G., & De Boer, H. (2020) *Soft skills in education. Putting the evidence in perspective*. Cham, Switzerland: Springer Nature

CV Bernard Veldkamp



Bernard Veldkamp is hoogleraar onderzoeksmethodologie en data-analyse. Hij is een expert in psychometrie, data science en geautomatiseerd adaptief testen. Zijn onderzoek richt zich op methoden voor dataverzameling en datagebruik voor het beoordelen van menselijk gedrag. Zijn missie is om psychometrie en data science te integreren om meten in de sociale wetenschappen te optimaliseren. Bernard promoveerde op het onderwerp geautomatiseerde testconstructie. Aanvankelijk richtte hij zich op het gebruik van multidimensionale modellen voor itemresponstheorie en hun implicaties voor testconstructie. Later richtte hij zich op de optimale samenstelling van itembanken, zowel voor geautomatiseerde adaptieve testen als voor lineaire testen. Toen er steeds meer data over menselijk gedrag beschikbaar kwamen, online, van sensoren en van social media, ging hij op zoek naar nieuwe methoden om deze gegevens te gebruiken bij het meten van menselijk gedrag. Hij verlegde zijn aandacht naar datamining en de combinatie van psychometrie en data science. Momenteel werkt hij aan het ontwikkelen van methoden voor het omgaan met multimodale gegevens (dat wil zeggen gegevens afkomstig uit verschillende bronnen) voor het meten van individueel en groepsgegedrag.

5 publicaties Bernard Veldkamp

Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K., & Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational research review*, 17, 50-62.

de Klerk, S., Veldkamp, B. P., & Eggen, T. J. (2015). Psychometric analysis of the performance data of simulation-based assessment: A systematic review and a Bayesian network example. *Computers & Education*, 85, 23-34.

Veldkamp, B. P., & Sluijter, C. (2019). *Theoretical and Practical Advances in Computer-based Educational Measurement*. Springer Nature.

Veldkamp, B. P., & Matteucci, M. (2013). Bayesian computerized adaptive testing. *Ensaio: Avaliação e Políticas Públicas em Educação*, 21(78), 57-82.

Van Der Kleij, F. M., Eggen, T. J., Timmers, C. F., & Veldkamp, B. P. (2012). Effects of feedback in a computer-based assessment for learning. *Computers & Education*, 58(1), 263-272.

CV Martina Meelissen

Dr. Martina Meelissen is verbonden aan de vakgroep Onderzoeksmethodologie, Meetmethoden en Data-analyse van de Faculty of Behavioural, Management and Social sciences van de Universiteit Twente. Momenteel is ze National Program Manager en penvoerder van het consortium (Universiteit Twente, Expertisecentrum Nederlands en KBA Nijmegen) dat verantwoordelijk is voor de Nederlandse bijdrage aan PISA 2018 en 2022 (Programme for International Student Assessment). Sinds 2003 is ze de National Research Coordinator van TIMSS (Trends in International Mathematics and Science Study). In die hoedanigheid is ze recentelijk verantwoordelijk geweest voor de uitvoering het Nederlandse aandeel in eTIMSS 2019 (toetsafname met tablets in groep 6) en voor (een deel van) de dataverzameling voor Peil.Rekenen-Wiskunde 2019. Ze maakt sinds 2002 deel uit van de internationale Questionnaire Item Review Committee van TIMSS, een expertgroep verantwoordelijk voor de contextvragenlijsten in TIMSS.

5 publicaties Martina Meelissen

Hamhuis, E., Glas, C., & Meelissen, M. (2020). Tablet assessment in primary education: Are there performance differences between TIMSS' paper-and-pencil test and tablet test among Dutch grade-four students?. *British journal of educational technology*, 51, 2340-2358. doi:10.1111/bjet.12914

Punter, R. A., Meelissen, M. R. M., & Glas, C. A. W. (2016). Gender differences in computer and information literacy: An exploration of the performances of girls and boys in ICILS 2013. *European Educational Research Journal (special issue)*, 1-19, doi: 10.1177/1474904116672468.

Mejía-Rodríguez, A. M., Luyten, H., & Meelissen, M. R. M. (2020). Gender Differences in mathematics self-concept across the world: An exploration of student and parent data of TIMSS 2015. *International Journal of Science and Mathematics Education* <https://doi.org/10.1007/s10763-020-10100-x>

Drent, M., Meelissen, M. R. M., & Kleij, F. M. van der (2013). The contribution of TIMSS to the link between school and classroom factors and student achievement. *Journal of Curriculum Studies*, 45(2), 198-224, <https://doi.org/10.1080/00220272.2012.727872>.

Gubbels, J. C. G., van Langen, A. M. L., Maassen, N. A. M., Meelissen, M. R. M., Aalders, P., Dood, C. M. M., & Wolbers, M. H. J. (2019). Resultaten PISA-2018 in vogelvlucht

Oberon

Postbus 1423, 3500 BK Utrecht

t 030 230 60 90 | f 030 230 60 80

info@oberon.eu | www.oberon.eu

Utrecht, april 2021

In opdracht van Ministerie van Onderwijs, Cultuur en Wetenschap